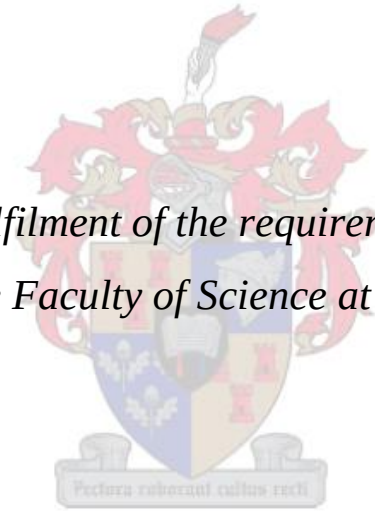
Indole- and diaryl ether-derived non-nucleoside reverse transcriptase inhibitors

by

Petronella Magdalena Wessels



Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University

Supervisor: Prof. Willem A. L. van Otterlo

Co-supervisor: Dr Stephen C. Pelly

Department of Chemistry and Polymer Science

Faculty of Natural Science

December 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2017

Abstract

Since the discovery of the Human Immunodeficiency Virus (HIV) as the cause of Acquired Immune Deficiency Syndrome (AIDS) in the early 1980s, no cure has been found. This disease has thus claimed the lives of millions of people and has gained the reputation as one of the worst pandemics in human history. There has, however, been considerable success in reducing the viral load of infected individuals through the use of combination therapy, known as highly active antiretroviral therapy (HAART). A main target of this drug regimen is the reverse transcriptase (RT) enzyme, which is inhibited by two of the three drug classes typically employed by HAART, namely nucleoside RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (NNRTIs). However, as the viral population diversifies through mutations introduced during viral replication, drug resistant viral strains emerge and limit the use of currently effective drugs. These resistant strains thus require new small molecule inhibitors to keep them in check.

In light of the need for new therapeutic agents to continuously enter the drug development pipeline, we set out to improve upon current NNRTIs through the lead optimization of an indole core structure and the design of a diaryl ether NNRTI using molecular modelling.

The first focus was to overcome the acid-catalysed indole-mediated degradation of the current lead compound, ethyl 5-chloro-3-[(methylthio)(phenyl)methyl]-1*H*-indole-2-carboxylate (designed in an earlier research project), thereby making the small molecule suitable for oral intake. We successfully synthesised twelve acid stable derivatives of this lead compound by replacing the thiomethyl group with an alkyl or alkene chain in the benzylic position. Four of these newly synthesised derivatives proved to be more potent (0.010 – 0.019 μM) than the lead compound (0.039 μM) against wild-type HIV-1, while six of these compounds were more potent than nevirapine (0.091 μM) under the same conditions.

The next study involved modifications to the 2-position of a lead indole based NNRTI, ethyl 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylate, through the introduction of various pyridinyl groups to determine the effects on efficacy. Three of the five resultant compounds performed better against wild-type HIV-1 (0.030 – 0.043 μM) than the lead compound, with an ethyl ester in the 2-position (0.060 μM), and two of these three compounds remained potent against HIV-1 viruses harbouring the prevalent NNRTI resistant mutations V106M, Y188C/H, G190A and K103N.

Due to the positive results obtained when an indole was used as a core structure in NNRTIs, ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate was chosen for the incorporation of electrophilic warheads as amides with the aim of synthesising the first irreversible NNRTIs. Three proof-of-

concept compounds were successfully synthesised with electrophilic warheads positioned to interact with the thiol of cysteine in the Y181C mutant strain of the virus. These compounds, however, proved to be toxic and thus ineffective as NNRTIs.

Finally, molecular modelling was used to design flexible diaryl ether NNRTIs with the ability to form desirable interactions with residues in the binding pocket. We sought to achieve an improved resistance profile through π-π stacking interactions with conserved amino acid residues Trp229 and Tyr318 and additional interactions with Tyr188 and Lys101. Three proof-of-concept compounds were then synthesised, of which 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile proved to be highly effective with an inhibition activity (IC_{50} value) of 5 nM against wild-type HIV-1.

Uittreksel

Sedert verworwe immuniteitsgebreksindroom (VIGS) in die vroeë 1980s as oorsaak van die menslike immuniteitsvirus (MIV) geïdentifiseer is, is daar nog geen geneesmiddel beskikbaar teen MIV nie. As gevolg hiervan het hierdie siekte al die lewens van miljoene mense geëis en die reputasie as een van die ernstigste pandemies in menslike geskiedenis gekry. Ten spyte van bogenoemde is daar al heelwat sukses behaal in terme van ´n verlaging van die virale vlakke in besmette individue deur gebruik te maak van kombinasie terapie naamlik hoogsaktiewe antiretrovirale terapie (HAART). Twee uit die drie geneesmiddels wat oor die algemeen deur HAART gebruik word, naamlik nukleosied trutranskriptase inhibeerders (NRTIs) en nie-nukleosied trutranskriptase inhibeerders (NNRTIs), teiken die trutranskriptase ensiem deur inhibisie. Mutasies gedurende virale replikasie lei egter tot die diversivering van die virale bevolking en so ontwikkel weerstandige stamme wat dan die gebruik van hierdie huidige effektiewe middele belemmer. Weerstandige stamme benodig dus nuwe klein molekule inhibeerders om hulle in toom te hou.

As gevolg van die konstante aanvraag na nuwe inhibeerders in die produksiepyplyn van geneesmiddels, is daar gepoog om te verbeter op die huidige beskikbare NNRTIs. Die indool-gebaseerde primêre-struktuur is gebruik, as ook die ontwerp van ´n di-ariel eter NNRTI deur die gebruik van molekulêre modelering.

Die fokus was aanvanklik op die uitskakeling van suurgekataliseerde indool-bemiddelde degredasie van die huidige primêre-struktuur, etiel 5-chloor-3-[(metieltio)(feniel)metiel]-1*H*-indool-2-karboksilaat (gesintetiseer in ´n vorige navorsingsprojek). Hierdeur sal die klein molekule per mond geneem kan word. Twaalf suurstabiele variasies van die primêre-struktuur is daarna suksesvol gesintetiseer deur die tiometielgroep met ´n alkiel- of alkeenketting in die bensiliese posisie te vervang. Vier van hierdie nuut gesintetiseerde afgeleide molekules was meer aktief (0.010 – 0.019 µM) as die primêre-struktuur (0.039 µM) teen wilde-tipe MIV-1 bevind, terwyl ses uit die twaalf molekules meer aktief as nevirapine (0.091 µM) onder diesefde kondisies bevind is.

Die volgende ondersoek het veranderinge tot die 2-posisie van die primêre-struktuur indool NNRTI, etiel 5-chloor-3-[(3,5-dimetielfeniel)(metieltio)metiel]-1*H*-indool-2-karboksilaat, behels. Verskeie piridiniel groepe is in hierdie posisie geplaas om die effek van die veranderinge op die doeltreffendheid teen MIV-1 te bepaal. Drie uit die vyf gesintetiseerde molekules was meer aktief bevind teen wilde-tipe MIV-1 (0.030 – 0.043 µM) as die primêre-middel (0.060 µM), wat ´n etiel ester in die 2-posisie het. Twee uit dié drie molekules het ook aktief getoets teen MIV-1 virusse met algemene NNRTI weerstandige mutasies V106M, Y188C/H, G190A en K103N.

As gevolg van die positiewe resultate wat deur die indool-gebasseerde NNRTIs verkry is, kon etiel 3-(1-aminopropiel)-5-chloor-1*H*-indool-2-karboksilaat gebruik word vir die inkorporering van elektrofiliese lokvalle as amiede met die doel om die eerste onomkeerbare NNRTIs te sintetiseer. Drie bewys-van-konsep molekules is suksesvol gesintetiseer met die elektrofiliese lokvalle nagenoeg aan die tiol groep van sisteïen, in die Y181C mutante vorm van die virus, gepossisioneer om ´n binding te vorm. Daar was ongelukkig bevind dat hierdie verbindings giftig vir lewende selle is en dus onbruikbaar as NNRTIs is.

Laastens is daar van molekulêre modelering gebruik gemaak om buigbare di-ariel eter NNRTIs te ontwerp wat die vermoë kan besit om gewensde interaksies met aminosuurresidue in die bindingsholte te hê. Ons het gemik om ´n verbeterde weerstands profiel te verkry deur middel van π-π stapelings interaksies met amienosuurresidue, Trp229 en Tyr318, wat nie maklik mutasie ondergaan nie, asook interaksies met Tyr188 en Lys101. Drie bewys-van-konsep molekules is toe gesintetiseer, waarvan 3-(3-amino-2-methoxyfenoxy)-5-chloorbenzonitriel hoogs aktief was met ´n inhibisie aktiwiteit van 5 nM teen wilde-tipe MIV-1.

Acknowledgements

Personal Acknowledgements

The journey I took from making plant-based “potions” in our back yard, through minor explosions during my undergraduate studies, to knowing what the buckminsterfullerene is and doing synthesis under inert conditions, seems like a story written by J. K. Rowling. I am not saying it was all a fairytale, but a few people with special powers made the impossible possible.

Firstly, I would like to thank my mom for helping me believe in the things that we cannot see and my dad, the psychiatrist, who says that strange is the new normal. Thank you for your endless support and love, for late nigh tea, for sweet treats (which my mom likes to call “hupstootjies”) and the many many phone calls every day. One of my favourite memories of my MSc is the day when my parents came to visit me in the lab and took so many photos of me in my lab coat and goggles that I felt like a celebrity. I don’t have the words to express my appreciation.

Next, I would like to thank my two grannies Lucille and Leen for believing that I am extraordinary. It has been said that people become what they are told they are. Ouma Leen, baie dankie dat Ouma altyd vir my sê “Lanatjie, jy kan alles doen” en Ouma Lucy, dat Ouma by die ouetehuistannies spog en sê “my poppie aard na haar slim oupa”.

To my boyfriend, Rick, thank you for being extremely supportive and encouraging from the day I met you in the crystallography laboratory, as a confused honours student, three years ago. Thank you for standing by my side through tears and laughter and for never failing to make me smile after a long day. After this thesis, I think you deserve a holiday more than I do.

I would also like to thank my two best friends Nicolaas and Natja. Nicolaas, thank you for our Friday afternoon lunch dates, for making me laugh and for keeping me sane. Natja, thank you for distracting me with stories about boys and for making yourself comfortable in my livingroom while I wrote my thesis. Few people are as lucky to have their very own cheering squad as I am.

To my supervisor, Prof van Otterlo, also known as the big friendly giant (BFG), I could not have completed this MSc without you. Thank you for your guidance and support, the many valuable discussions and wonderful opportunities I have had during this project. Thank you for always having time for me. I am going to let you into a secret: sometimes I just knock on your door and pop in, because I like seeing your friendly face behind the massive pile of papers on your desk. You are much more than a supervisor to your chemistry kids. You inspire and encourage us, you allow us to

be creative, have fun and explore and when we make mistakes, you tell us “that is the best way to learn” and you turn all negativity up-side-down.

Dr Pelly, thank you for your guidance and advice during this project and for making time for Skype meetings, emails and many valuable discussions. I am very fortunate to have had one local and one international supervisor who put a lot of time and effort into their students.

Prof Green, I thoroughly enjoyed sharing a lab with you. You have so many stories to tell, advice to give and knowledge to share. I regard myself as very privileged to have met you. Thank you for your help when it came to synthesis problems and interpreting complex NMR spectra. The lab would have been dull without your warm smile.

Dr Arnott, your interest and enthusiasm for chemistry is contagious. I really appreciate all your help, not only during my MSc, but throughout my studies, with mechanisms, synthetic problems and pretty much anything to do with computers. You are always willing to help and your door is always open for any curious or confused student.

Dr Blackie, whenever I have to give a presentation I always look for your friendly face in the audience and think “Dr Blackie is on my side” and then I feel like I could give an inauguration speech. Thank you for being so supportive. You may not know this, but they often quote you in church, which is when I whisper proudly to the person next to me “I know her” and think, “she must have inspired many students by pursuing your passions”.

Earlier this year I had the privilege of meeting Dr Adri Basson in person. I was lost in one of the biggest buildings on the Wits campus, where the different floors were given seemingly arbitrary numbers and the only people around were sitting in the cafeteria at the entrance of the building, possibly because they gave up with the lifts a while ago. I nearly gave you a hug when you told me you were Adri. Thank you for all the hours you put in doing our biological testing for this project, for working extremely fast to satisfy our curiosity and for answering all my questions with regards to methods and assays. Adri, jy is ´n staatmaker.

On the analysis side, I would like to thank Dr Jaco Brand and Ms Elsa Malherbe for running the NMR unit so efficiently and for always being prepared to help the students with anything from computer problems to spectral analysis. Your kindness and smiles are very much appreciated. For mass spectral analysis I would like to thank Dr Marietjie Stander for her hard work and invaluable contribution to our publications and theses.

Finally, I would like to thank the exceptional GOMOC team for help, support and lots of fun. A special thanks to the girls in Lab 2020/21, Shivvy, Leandi, Nicole and Tanya, for advice, support, friendship and for keeping my mistakes secret.

From this very long list of acknowledgements, it is clear that I have the most amazing people in my life. A final very big thank you! to each of you.

Acknowledgements for funding

Without funding this MSc project would not have been possible. I would therefore like to thank the National Research Foundation (NRF) for their generous contribution in the form of the grant-holders linked bursary and for their efforts in promoting research in South Africa.

Make sure everything you do is so completely crazy it’s unbelievable ~ Roald Dahl, Matilda

Table of Contents

Chapter 1: The Discovery and History of HIV as the Cause of AIDS, the Mode of Infection of the Virus and the Search for Treatments

1.1 Introduction

Acquired Immune Deficiency Syndrome (AIDS) is a disease caused by the Human Immunodeficiency Virus (HIV) which, with no viable cure, is known to be one of the worst pandemics in human history.[1,2] In 2016, this disease was responsible for an estimate of 1 million deaths globally, while 36.7 million people are living with HIV.[3] This disease has far reaching implications in terms of economic, social and political stability, especially in developing countries where medical resources are limited.[1,2]

In South Africa alone, it is estimated that 110 000 people died in 2016 as a result of AIDS and 7.1 million adults and children are HIV positive.[4] It is predicted that AIDS related deaths will reduce the South African adult population by 7.8 million people by 2025. This will have grave implications for population growth, labour supply, labour productivity, education, healthcare and social welfare and in turn the economy of our country.[5,6] It is therefore no longer only a healthcare crisis, but affects all sectors.[6] KwaZulu-Natal (KZN), the epicentre of South Africa's HIV pandemic, highlights the seriousness of the epidemic with 26.4% of the working age population being HIV positive, compared to 15.9% in the rest of the country. By 2025, the economy of KZN is predicted to be 43% smaller as a direct consequence of HIV/AIDS, while the rest of the country is similarly affected with a 37% smaller economy.[5] Since our country has the largest number of HIV infections in the world, combating the AIDS epidemic has thus become a priority.[7]

1.2 A brief history

1.2.1 Discovery

The discovery of HIV as the cause of AIDS was a serendipitous event, starting with the detection of cancer-causing retroviruses in animals in the 1960s.[8] This encouraged the development of the U.S. Virus Cancer Program, which aimed to identify tumour-causing viruses, especially retroviruses, in humans.[8] By the late 1970s, however, it was believed that viruses were not responsible for any human cancers and that human retroviruses did not exist.[8–10] Furthermore, serious epidemic diseases were thought to be limited to developing nations.[8,9] Under this false sense of security, the U.S. Virus Cancer Program was terminated and many microbiology departments shifted their focus.[8] Robert C. Gallo and Luc Montagnier, together with their co-workers, however, made tumour-causing

human retroviruses their primary research objective, their groups focusing on breast cancers and leukemias.[8–10] The discovery of reverse transcriptase (RT) in 1970 by Temin and Baltimor, proved to be central in the search for retroviruses.[8–10] Since RT is present in all retroviruses, assays to detect RT could be used to search for retroviruses.[8,9] Success was obtained when RT from human lymphocytic leukaemia cells were purified in 1972.[8] All that remained was to detect the human retrovirus producing the RT. Human T-cell leukaemia viruses, types 1 and 2 (HTLV-1 and HTLV-2), were consequently discovered and isolated by Gallo and Montagnier in 1979 and 1982 respectively, marking the first identification of a human retrovirus.[2,8–10]

The link between retroviruses and AIDS was made when animal studies revealed that HTLV could cause AIDS-like symptoms and risk factors.[8,9] It was known, for example, that variants of feline leukaemia retrovirus could cause immune disorders by targeting CD4+T cells and that the mode of infection of HTLV-1 includes blood exposure, sexual contact and mother to child transmission during childbirth.[8,9] A new retrovirus, today known as HIV, was first isolated in 1983 by Montagnier and Chermann from blood cell cultures of a man who suffered from lymphadenopathy, a syndrome that was considered to be a precursor of AIDS.[8,9]

Advancements, such as the ability to grow human T cells (1976), the development of sensitive and specific assays for retroviruses (1970s), the development of the HIV blood test that allowed large-scale screening and verification of HIV (1983, 1984) and the discovery of zidovudine (AZT) as an inhibitor of the HIV reverse transcriptase (1985, 1986), pointed toward the rapid eradication of the disease.[10] Scientists with experience in retrovirology, however, knew that unless a vaccine was developed soon, the infection may become permanent in the population as is the case with retroviruses in many species.[11] The use of triple drug therapy in 1995 greatly enhanced the quality of life of HIV positive people and mortality approached a normal life expectancy.[11] Currently, mutant forms of the virus developing under drug selective pressure has resulted in a decrease in the effectiveness of this therapy.[12,13] Continued research towards the development of new antiviral drugs is thus of great importance.[10,13]

1.2.2 Evolution

The evolution of the diversity in HIV-1, as seen from samples around the world, gave clues as to the origin of the virus.[14] The HIV-1 strain can be subdivided into three distinct groups, namely the rare N and O groups, restricted to Cameroon and neighbouring countries, and the highly prevalent M group, with its epicentre of diversity situated in Kinshasa, DRC, Africa.[14,15] This extent of accumulated diversity indicates that the virus has circulated within human populations for many years prior to its recognition in 1981.[14] At the time of its discovery, the closest known relative to HIV-1 was visna, a sheep Lentivirus. Other lentiviruses were soon identified in non-human primates, as well as in

humans (HIV-2). The viruses from non-human primates were called simian immunodeficiency viruses (SIVs) and have thus far only been found to naturally infect primates in sub-Saharan Africa. Multiple strains of SIVs found in a single species form a monophyletic clade, indicative of intraspecific transmissions. In the same manner, primate viruses, including HIV, form a distinct clade within lentiviruses. This indicates that humans acquired their infections from other primates. The evolutionary origin of HIV was therefore expected to be from a non-human primate.[16]

In 1989, a SIV closely related to HIV-2 was found in the sooty mangabey monkey occurring naturally in west Africa, and was termed SIVsmm.[17] A virus closely related to SIVsmm was then found in captive macaques and was noted to cause severe AIDS-like symptoms.[18] It soon became apparent that SIVsmm had been unintentionally transmitted to macaques in captivity.[19] Because the HIV-2 strain can be divided into various groups which are interspersed among the SIVsmm lineages, the sooty mangabey were identified as the source of HIV-2. Multiple mangabey-to-human cross-species jumps explain the interspersion.[19–21]

In the same year, 1989, a virus closely related to HIV-1 was found in two captive chimpanzees in Gabon and was named SIVcpz.[22,23] After testing numerous other chimpanzees over the period of 10 years, only a single case of the virus was found in a chimpanzee in Belgium that was illegally imported from Kinshasa.[24] It appeared that SIVcpz was very rare in chimpanzees and that both chimpanzees and humans obtained the infection from some other source(s). A fourth example of SIVcpz in chimpanzees was obtained in 1999 in a wild-caught individual of unknown origin imported to the USA.[16] Mitochondrial DNA analysis of the four chimpanzees revealed that the two chimpanzees from Gabon and the one from the USA belonged to the *P. t. troglodyte* subspecies, while the other chimpanzee in Belguim was a *P. t. schweinfurthii.*[25] The majority of the chimpanzees tested belonged to the *P. t. versus* subspecies where the infection was absent, resulting in deceiving data.[26,27] SIVcpz from *P. t. troglodytes* formed a monophyletic cluster with all HIV-1 strains, implicating one specific chimpanzee subspecies as the source of the human virus.[16,25,26] From studies conducted on two habituated communities of chimpanzees at Gombe National Park in Tanzania, it was found that SIV causes similar symptoms in chimpanzees to AIDS in humans. As in infected humans, significant depletion of $CD4^+$ T-cells were observed in infected chimpanzees.[28]

1.3 Mode of infection

The two major types of the HI virus, HIV-1 and HIV-2, are believed to have originated from the transmission of SIVcpz and SIVsm from two different primate species to humans.[11] It was hypothesised that the outbreak of the virus in humans resulted from cutaneous or mucous membrane exposure to infected animal blood. Direct exposure to animal blood and secretions as a result of

hunting, butchering or the consumption of uncooked, contaminated meat have been suggested as likely means of transmission.[21] Human-human transmissions occur most commonly through sexual behaviours or needle sharing. When bodily fluids, such as blood, semen, pre-seminal fluid, rectal fluids, vaginal fluids or breast milk, from a person who has HIV come into contact with the mucous membrane or damaged tissue or is directly injected into the bloodstream, the virus may be transmitted from person to person.[29] To account for the appearance of AIDS as an epidemic in the 20th century, and not before, social, economic and behavioural changes in human populations have been suggested. These include social disruption, enslavement, urbanisation, prostitution, and the use of non-sterilised needles for parenteral injections and vaccinations.[21]

Once the virus has entered the body of the host, the outward protruding viral envelope protein, called the docking glycoprotein (gp120), interacts with CD4 receptors on the immune system's helper T lymphocytes and macrophages (*Figure 1*, Binding).[12,30] The CD4 molecule was the first retroviral receptor identified and is the primary target of the HI virus.[31] Following this initial interaction is the binding of gp120 to co-receptors on the host cell surface. Two co-receptors, α- and β-chemokine, dubbed CXCR-4 and CCR5 respectively, initiate fusion of the lipid bilayers of the viral envelope and the host cell plasma membrane.[12,31] T-cells most commonly express CXCR4 rather than CCR5, while primary lymphocytes express both CXCR4 and CCR5, and macrophages express CCR5. Virus strains that preferentially use the CXCR4 co-receptor are called X4 viruses, while strains using the CCR5 co-receptor are called R5 viruses. Dual-tropic virus strains using both receptors are denoted R5X4 viruses.[31] Co-receptor binding results in conformational changes to the transmembrane glycoprotein (gp41) on the viral envelope so that it is anchored into the host cell membrane. This ultimately instigates the release of the viral core into the host cell's cytoplasm (*Figure 1*, Fusion).[12,31] Once in the cytoplasm, the viral core, containing genomic material, including RNA and enzymes such as reverse transcriptase (RT), integrase and protease, is converted into the reverse transcription complex (RTC) and then the pre-integration complex (PIC) (Uncoating, *Figure 1*, A).[12,31] A defining characteristic of retroviruses is their ability to convert their RNA genomes into double-stranded (ds) DNA (*Figure 1*, Reverse Transcription). Each viron holds two single-stranded (ss) copies of its RNA genome.[31] RT catalyses the synthesis of complimentary DNA strands forming RNA/DNA hybrids after which the RNA strands are cleaved by the RNase H domain of RT.[31,32] RT then catalyses the synthesis of dsDNA from the remaining ssDNA.[30] Viral proteins vpr, matrix protein (MA) or Integrase facilitates the movement of viral dsDNA through pores in the nuclear membrane (*Figure 1*, B). Inside the nucleus, HIV Integrase catalyses the integration of proviral dsDNA into the host's chromosome (*Figure 1*, integration).[12,31] The DNA can now either remain latent or can be translated into viral RNA and mRNA containing the genetic material to produce new viruses and the necessary viral proteins respectively (*Figure 1*, Translation). Translation of mRNA results in non-functional poly-proteins (Gag and GagPol) which can be cleaved by protease to form functional

proteins and enzymes for the assembly of the new virus (*Figure 1*, C).[2,12,31,] The immature virus then buds off the host cell and uses a piece of cell membrane to form its own viral membranes (*Figure 1*, Budding). The virus matures by processing its viral proteins to be able to infect new cells (*Figure 1*, Maturation). Finally, host cells weaken with the release of many viruses and ultimately die.[12]

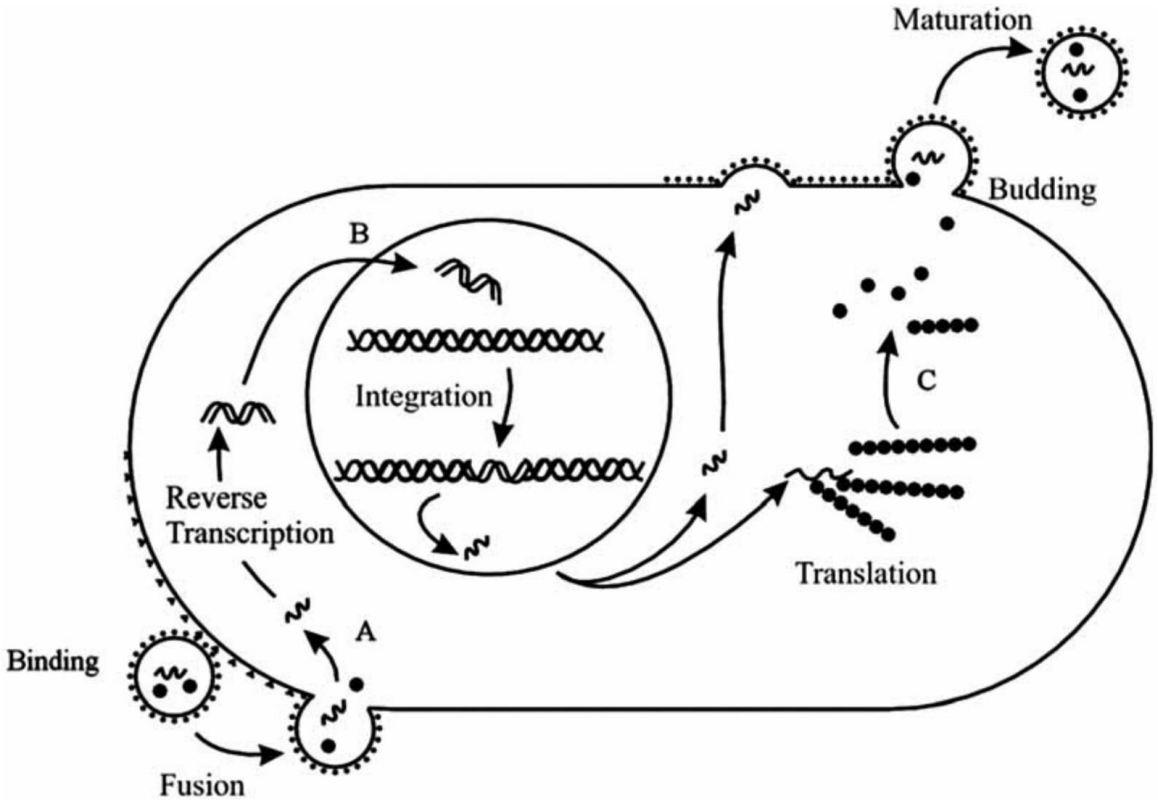


Figure 1: The HIV life cycle (obtained from "The Current Status of the NNRTI Family of Antiretrovirals Used in the HAART Regime Against HIV Infection").[12]

Following primary infection, the virus replicates rapidly resulting in a high viral load in the infected individual. At this stage of the disease, transmission from person to person is more likely.[33] A phase termed clinical latency now follows. When the CD8+T cells and antibodies of the immune system fight the virus, HIV levels in the blood decreases dramatically and CD4+T cell levels increase. Although the infected individual usually shows no symptoms of the virus during this stage, viral replication continues in lymphoid organs and the immune system slowly weakens. Monocytes, macrophages and some CD4+ cells, which survive HIV infection, act as reservoirs of the virus. Towards the end of this period, the CD4+ cell levels however, decrease and the viral load increases.[2,34] With the weakening and eventual death of cells of the human immune system, opportunistic infections, such as pneumonia or tuberculosis, result in the disease state known as AIDS (less than 200 CD4+ T cells per cubic millimetre of blood).[2,30] This disease is characterised with weight loss, minor mucocutaneous disease, severe bacterial infection, chronic diarrhoea and chronic fever. In the absence of treatment, the average survival time of Africans infected with HIV-1

is approximately 10 years, with the leading cause of morbidity being tuberculosis and severe bacterial infections.[35]

1.4 In search of a treatment

Before 1996, HIV-1 treatment consisted of the management of AIDS-related illnesses and prophylaxis against common opportunistic pathogens.[36] The turning point in HIV-1 treatment was marked by the release of Azidothymidine (AZT) as the first drug for the treatment of HIV/AIDS in 1985 (*Figure 2*).[11] Dideoxythymidine-type compounds were synthesised for use as anti-cancer agents, mimicking DNA nucleosides and thereby acting as chain terminators in DNA synthesis. AZT, synthesised in 1964 by Jerome P. Horwitz, was amongst these compounds, but showed no anti-cancer activity.[2,37] In 1974, Wolfram Ostertag, at the Max Planck Institute for Experimental Medicine, tested AZT against Friend Leukaemia Virus (FLV). After it was found to be a successful inhibitor in murine cell culture, a US pharmaceutical company, Burroughs Wellcome, bought the compound to conduct further animal tests on it.[2,38] It was ultimately shelved for the lack of human application, since retroviruses were not known to infect humans.[2] In 1984, after the discovery of HIV/AIDS, pharmaceutical companies started screening their compound libraries for molecules showing promise as leads against HIV.[2] After showing positive results when scanned for activity against retroviruses FLV and Harvey Sarcoma Virus (HaSV), AZT was sent for testing against HIV by Samuel Broder and Hiroaki Matsuya at the US National Cancer Institute in Bethesda, Maryland. Its *in vitro* activity was then confirmed at Duke University by Dani Bolognesi.[38] In 1985, 21 years after its synthesis, Joseph Rafuse became the first person to take AZT for treatment against HIV. After six weeks, his T-cell count had elevated significantly.[2]

The development of other RT and protease inhibitors (PI) and later combination drug regimens revolutionised the treatment of this disease (*Figure 2*).[36,39] Combination therapy known as highly active antiretroviral therapy (HAART), using three (or more) antiretroviral agents directed against at least two distinct molecular targets, dramatically supresses viral replication and reduces the plasma HIV-1 viral load to below detection limits.[36,39–41] The regeneration of a patient's immune system is measured by the increase in CD4+ T-lymphocytes.[39,41] Importantly, this three-antiretroviral-drug regimen decreases the probability of the virus selecting for clones with multiple mutations that could gain drug resistance.[36] The RT process is error prone with one mutation introduced for every 1000-10,000 nucleotides synthesised.[42–44] Because the HIV-1 genome is ~10,000 nucleotides in length, one to ten mutations may occur per viral genome with every replication cycle.[36] As the viral population diversifies, a particular mutation that confers a small survival advantage in the presence of one particular drug will emerge. As this variant strain continues to replicate, additional mutations harbouring greater resistance to an antiretroviral agent will accumulate in the population. Eventually,

a variant able to fully resist the drug will emerge and will become the dominant strain.[45] In this manner the generation of genetic diversity enables the virus to escape single drug therapy.[46]

Mathematical modelling, based on HIV-1 turnover rate and virus population size, suggests that any combination of drugs in which at least three mutations are required for drug resistance, should provide durable inhibition.[46–49] Trial and error in clinical trials established the basic principles for effective drug combinations.[36] Generally, these combinations consist of two nucleoside reverse transcriptase inhibitors (NRTIs) and either a non-nucleoside reverse transcriptase inhibitor (NNRTI) or a PI.[50] The United States and European Union has recommended the initiation of the HAART regimen in a patient when the CD4 cells in peripheral blood decline to 350 per cubic mm.[51] In poorer-resourced countries such as South Africa, drugs are only administered by the government when the CD4+T cell count drops below 200 per cubic mm.[2] With the correct adherence, HAART can suppress viral replication to the point where a patient can lead a normal life with a regular life expectancy.[36] The number of pills that must be taken daily has been reduced from more than 20 pills a day to only 1 in 2006.[52] HAART is however not a cure and maintaining therapy for a lifetime presents challenges.[36] Drug resistance may still evolve with nonadherence, poor drug tolerability and drug interactions with antiretroviral agents and other medications.[36] As the virus continues to evolve and escape even the most effective drug therapies, new HIV-1 treatments will be needed continuously.

Drug discovery has been revolutionised with high throughput compound screens using virus-specific replication or enzymatic-assays, optimization of inhibitors using lead compounds based on homologous enzymes or targets and rational drug design modelled on the structures of viral proteins.[36] This has enabled the development of six distinct classes of antiretroviral drugs classified according to their molecular mechanism and resistance profiles. These are: cell entry inhibitors [co-receptor inhibitors (CRIs: maraviroc) and fusion inhibitors (FIs: enfuvirtide)], nucleoside reverse transcriptase inhibitors (NRTIs: abacavir, didanosine, emtricitabine, lamivudine, stavudine, zalcitabine and zidovudine), nucleotide reverse transcriptase inhibitors (NtRTIs: tenofovir disoprovil fumarate), non-nucleoside reverse transcriptase inhibitors (NNRTIs: delavirdine, efavirenz, etravirine, nevirapine and rilpivirine), integrase inhibitors (INIs: dolutegravir, elvitegravir, raltegravir) and protease inhibitors (PIs: amprenavir, atazanavir, darunavir, fosamprenavir, indinavir, lopinavir, nelfinavir, ritonavir, saquinavir and tipranavir).[36,52]

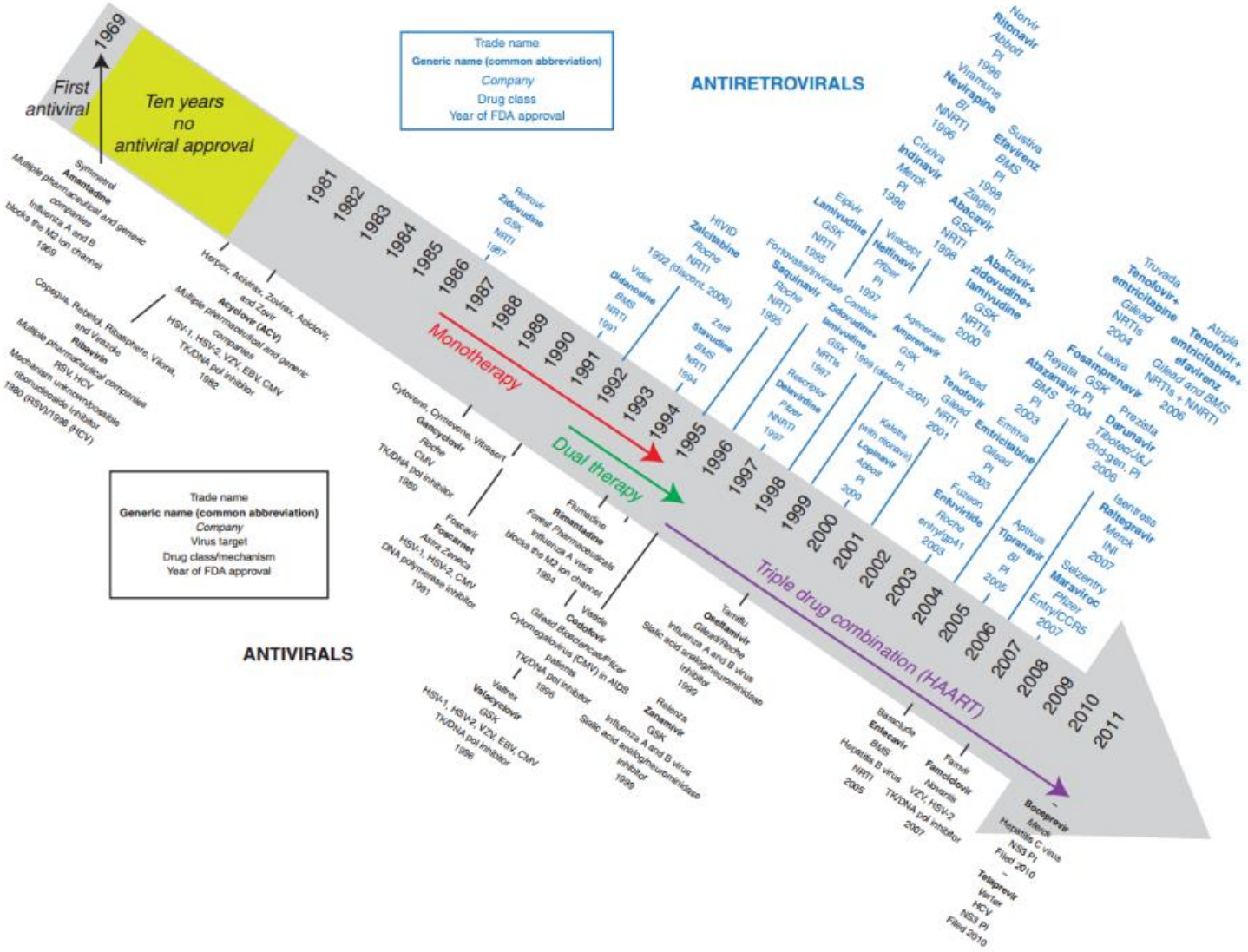


Figure 2: A timeline for the FDA approval of drugs currently in use as antiviral and antiretroviral agents (obtained from "HIV-1 Antiretroviral Drug Therapy").[36]

1.4.1 Inhibitors at the various stages of the viral life cycle

Chemical compounds able to selectively block HIV replication, without causing significant damage to the host, have been the subject of considerable research since the mode of infection of the virus has been elucidated. The HIV life cycle presents many opportunities for therapeutic intervention (*Figure 3*).[36] These are discussed and the successes are highlighted below. With HIV-1 being more infectious and thus more easily transmitted than HIV-2, this strain accounts for the vast majority of global HIV infections.[2] For this reason the discussion will mainly be focused on HIV-1.

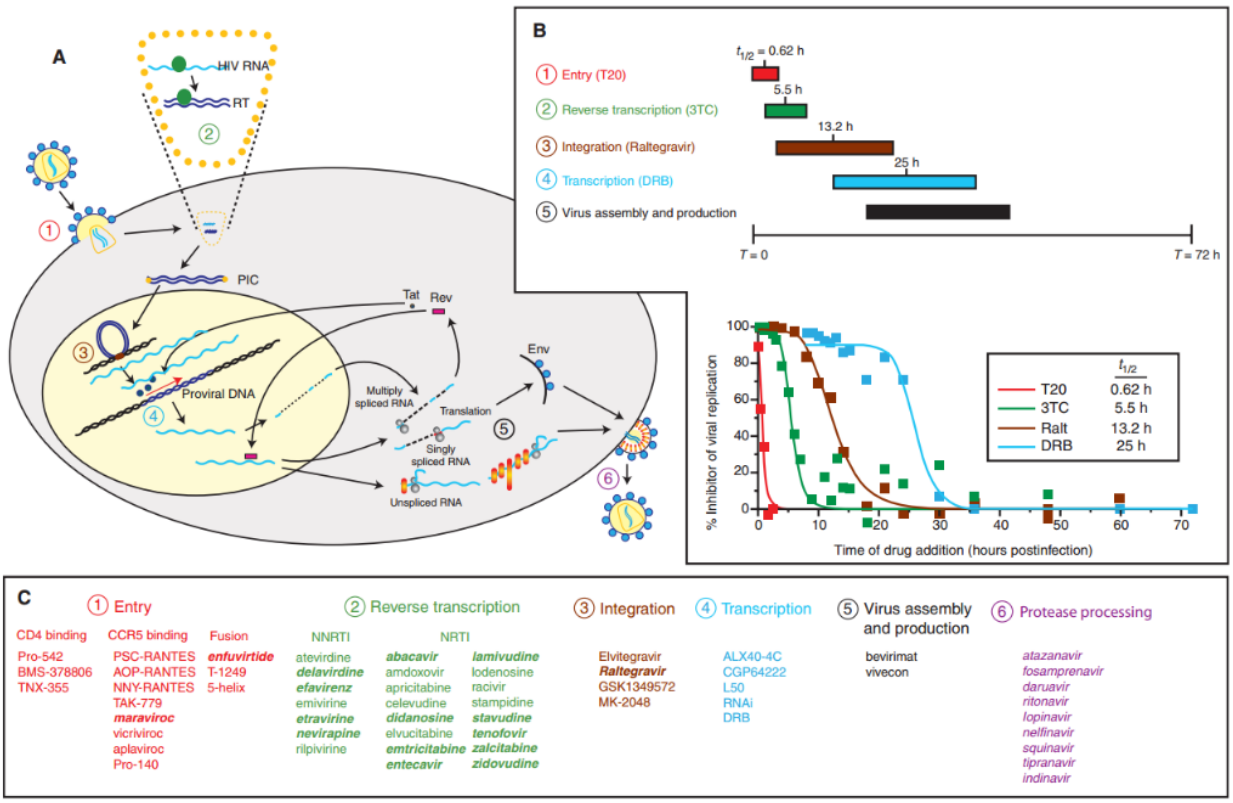


Figure 3: Steps in the HIV life cycle as current or potential targets for antiretroviral drugs (obtained from "HIV-1 Antiretroviral Drug Therapy").[36] (A) Steps in the HIV-1 life cycle blocked by approved inhibitors. (B) The time required for the completion of each of the steps in a single replication cycle. (C) Inhibitors in development (normal text); FDA approved drugs (italics/bold text).

1.4.1.1 Viral entry

The HIV-1 virus uses host proteins to drive a set of intricate events leading to membrane fusion and the release of the viral core into the cytoplasm.[36,49] Several classes of antiretroviral drugs target these viral entry proteins with the aim of inhibiting the first step in the HIV-1 replication cycle. These inhibitors include attachment inhibitors, chemokine receptor antagonists, and fusion inhibitors.[49]

1.4.1.1.1 CD4 attachment

Inhibitors targeting Gp120 and CD4, have showed clinical promise, but none have yet been approved for clinical use. Examples are the small molecule BMS-378806 and the CD4 antibody TNX-355.[53–55] BMS-378806 binds to the CD4 binding pocket on gp120, altering the conformation of the viral envelope protein in such a way that it cannot recognise CD4.[54] On the other hand, TNX-355 binds to CD4, inhibiting HIV-1 envelope docking without inhibiting CD4 function in the immune system.[55]

1.4.1.1.2 Coreceptor binding

The CCR5 coreceptor is used by the most commonly transmitted strains of HIV-1. Furthermore, the blocking of this coreceptor does not have a significant impact on human health as is indicated by the

normal immune function of individuals lacking CCR5. For this reason, CCR5 antagonists have been identified as suitable anti-HIV targets. Small-molecule CCR5 antagonists bind to a hydrophobic pocket within the transmembrane helices of CCR5, inducing and stabilizing conformational changes to other binding sites so that these are not recognised by CCR5 agonists or the HIV-1 envelope.[36,56,57] The CCR5 antagonist, CMPD167, tested as a microbicide in resus macaque, proved to be effective in inhibiting viral transmission.[58] Maraviroc (MVC, Selzentry®, *Figure 4*), a CCR5 antagonist, was approved for therapeutic use in humans in 2008. It is generally administered to treatment-experienced patients in whom other regimens have failed. Prior to prescription, the patient must be tested for the CCR5-tropic virus since the drug is ineffective against CXCR4- or dual-tropic viruses.[59] In addition, clinical trials with a second CCR5 antagonist, Vicriviroc, continues.

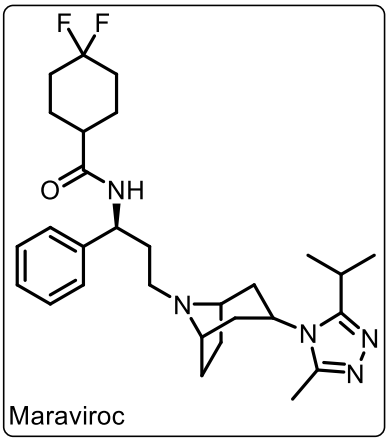


Figure 4: The CCR5 antagonist, maraviroc.

1.4.1.1.3 6-Helix bundle formation/membrane fusion

The fusion active conformation of gp41 is a six-helix bundle in which three helices form an interior, trimeric coiled coil onto which the other three antiparallel C helices pack.[60,61] Fusion of the virus to the host cell is promoted by the interaction of the two homologous domains in the viral gp41 protein.[36] When a heterologous protein mimics one of these domains, it can bind and disrupt the intramolecular interactions of the virus protein, thus blocking fusion with the host cell membrane.[36] Currently only one FI, enfuvirtide (ENF, T-20, Fuzeon®, *Figure 5*), is available for clinical use. As a synthetic polypeptide with 36 amino acids, this drug is not orally available and must be administered intravenously twice daily, making long-term use uncomfortable and problematic. Enfuvirtide is mainly used when standard combination regimens fail.[52,61,62]

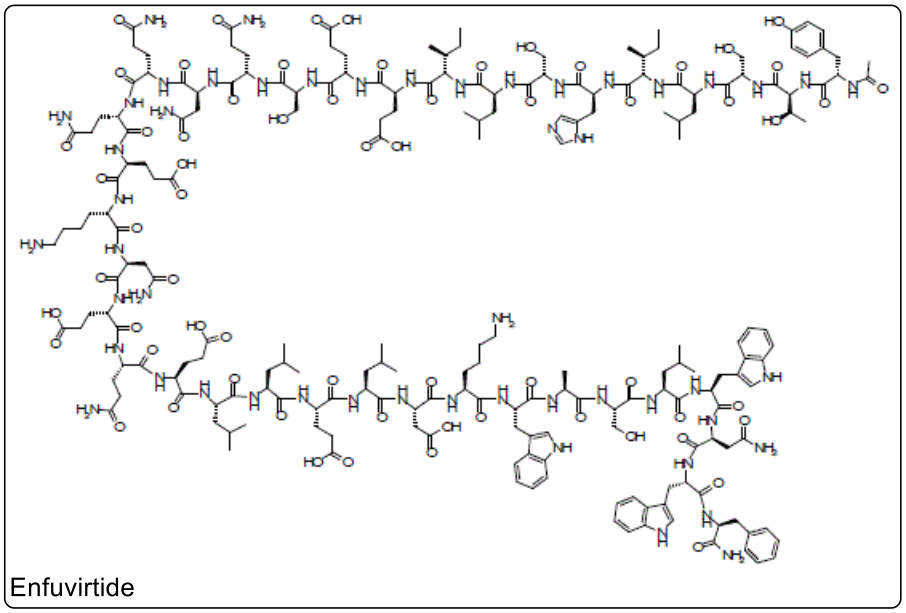


Figure 5: The fusion inhibitor, enfuvirtide.

1.4.1.2 Reverse transcription

RT was the first HIV-1 enzyme targeted in antiretroviral drug discovery with 12 licensed agents falling into the classes NRTIs, NtRTIs and NNRTIs. This accounts for nearly half of all approved antiretroviral drugs. RT was identified as an attractive therapeutic target, not only because it is essential for HIV replication, but because it is not involved in host cell metabolism.[63] The drugs affect the DNA polymerisation activity of the enzyme and block the generation of full-length viral DNA.[36]

1.4.1.2.1 Nucleoside/Nucleotide reverse transcriptase inhibitors

NRTIs and NtRTIs are analogues of naturally-occurring dideoxyribonucleosides (ddN). They are administered to patients as prodrugs requiring host cell entry and phosphorylation by cellular kinases to exhibit antiviral activity.[52,64] In the triphosphate form (ddNTP), they serve as competitive inhibitors, preventing the incorporation of the normal deoxynucleoside triphosphate (dNTP) substrate into the newly synthesised DNA chain.[38,52,65–68] The phosphorylation of the 3'-hydroxyl group at the sugar moiety of the drug prevents further chain elongation thereby inhibiting the production of either the (–) or (+) strands of the HIV-proviral DNA.[36,52,63] Seven NRTIs have been FDA approved for clinical use: abacavir (ABC, Ziagen®), didanosine (ddI, Videx®), emtricitabine ((-)FTC, Emtriva®), lamivudine (3TC, Epivir®), stavudine (d4T, Zerit®), zalcitabine (ddC, Hivid®), and zidovudine (AZT, Retrovir®) (*Figure 6*).[36]

NtRTIs are unique from NRTIs in that NRTIs require three phosphate groups to be added before they can become incorporated by RT, while NtRTIs already have one phosphate group in place and only require an additional two to be added. Since the first phosphorylation is the rate limiting step, the phosphorylation of NtRTIs, with one phosphate group already in place, leads to a faster

phosphorylation process to the triphosphate active drug. Tenofovir disoproxil fumarate (TDF, Viread®, *Figure 6*), an NtRTI, has become one of the most frequently prescribed drugs for the treatment of HIV.[52] Recently, a new pro-drug of tenofovir, namely tenofovir alafenamide (TAF, Vemlidy®, *Figure 6*), had been described as successor of TDF due to the 30-fold lower dose required to achieve the same potency. Three combination therapies with TAF have so far been approved for clinical use and are marketed as Genvoya®, Descovy® and Odefsey®.[69,70] Furthermore, Wyatt and Baeten hypothesized that TAF may be the first stepping stone toward regimens design for lifelong use.[71]

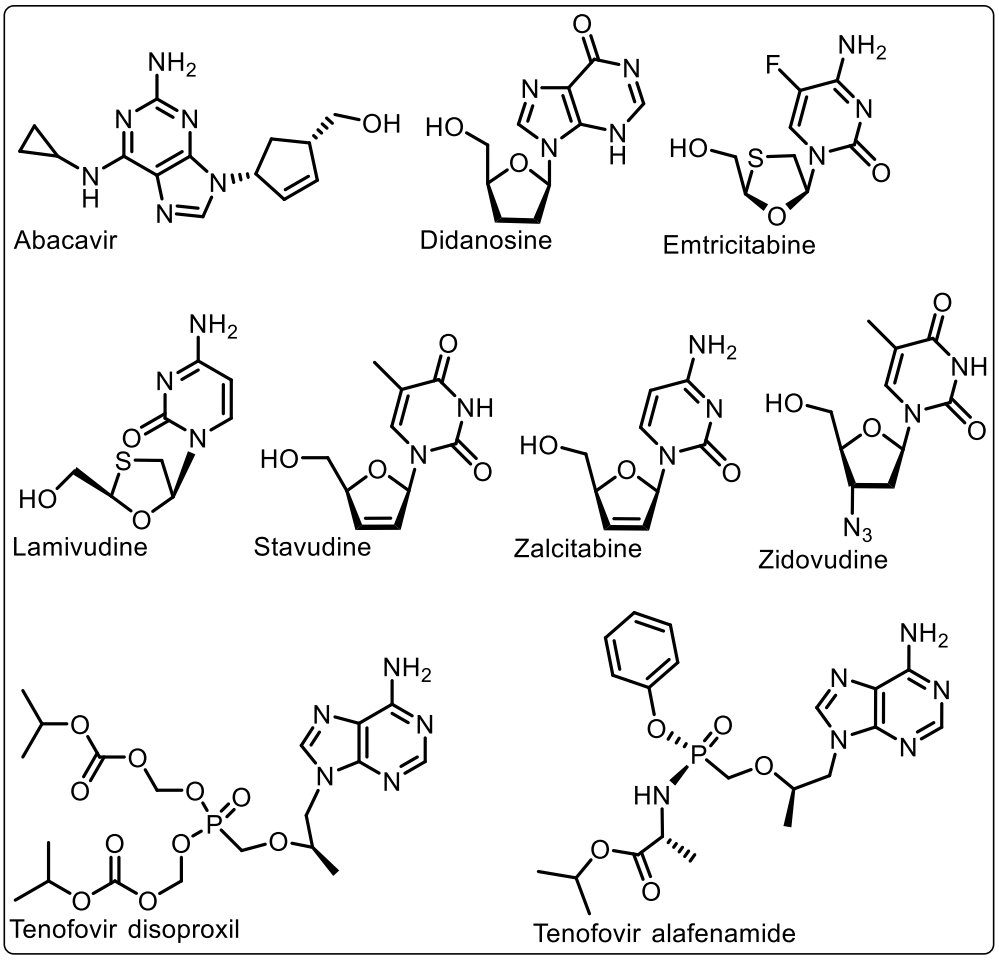


Figure 6: Seven FDA approved nucleoside reverse transcriptase inhibitors, abacavir, didanosine, emtricitabine, lamivudine, stavudine, zalcitabine and zidovudine and two nucleotide reverse transcriptase inhibitor, tenofovir disoproxil and tenofovir alafenamide.

1.4.1.2.2 Non-nucleoside reverse transcriptase inhibitors

NNRTIs inhibit the HIV-1 RT enzyme by binding to an allosteric, hydrophobic pocket, located approximately 10 Å from the catalytic site.[63,72,73] This causes a change to the spatial conformation of the substrate-binding site which in turn reduces polymerase activity.[63,73,74] HEPT (i.e. emivirine) and TIBO (i.e. tivirapine) derivatives were the first classes of antiretroviral compounds recognised as non-nucleoside inhibitors and paved the way for numerous potent inhibitors, namely delavirdine

(Rescriptor®), efavirenz (Sustiva®, Stocrin®), etravirine (Intelence®), nevirapine (Viramune®), and rilpivirine (Edurant®) (*Figure 7*).[50,52,73,75,76]

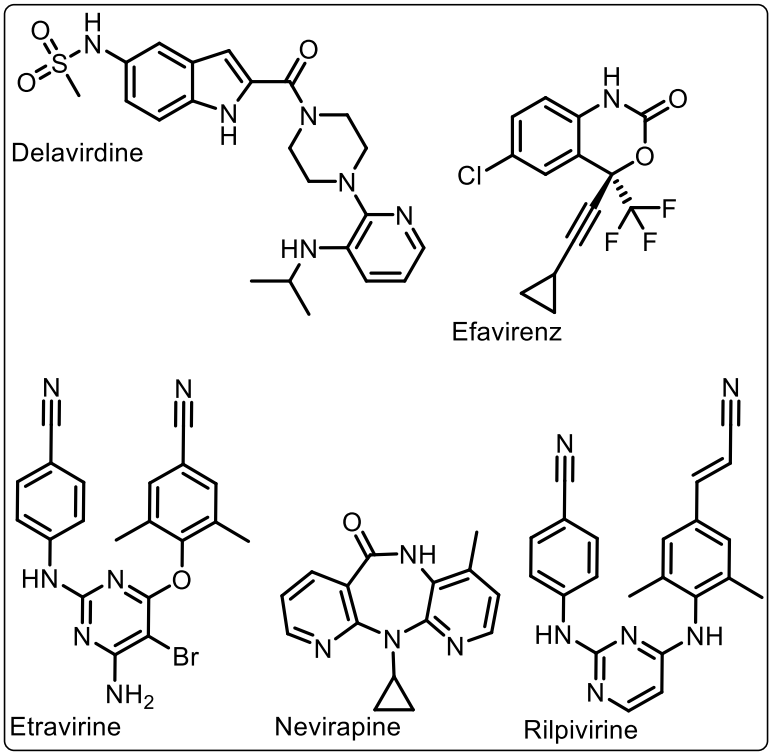


Figure 7: Non-nucleoside reverse transcriptase inhibitors currently on the market.

1.4.1.3 Integration

The most recent class of approved antiretroviral drugs, integrase inhibitors (INIs or InSTIs), target one of the three viral enzymes essential for viral replication.[2] The enzyme integrase catalyses 3' end processing and viral DNA strand transfer. The approved integrase inhibitors target the strand transfer reaction and block integration of HIV-1 DNA into the host cell DNA.[77–80] An INI inhibitor must be designed with a metal-binding pharmacophore which is able to interact with two magnesium or manganese ion cofactors in the active site and a hydrophobic group for viral DNA and the integrase-viral DNA complex interaction.[2,61,80,81] The first integrase inhibitor, Raltegravir (RAL, Isentress®), gained FDA approval in 2007.[61,80] Other integrase inhibitors include Dolutegravir (Tivicay®) and Elvitegravir (EVG, Vitekta®) (*Figure 8*).[82,83]

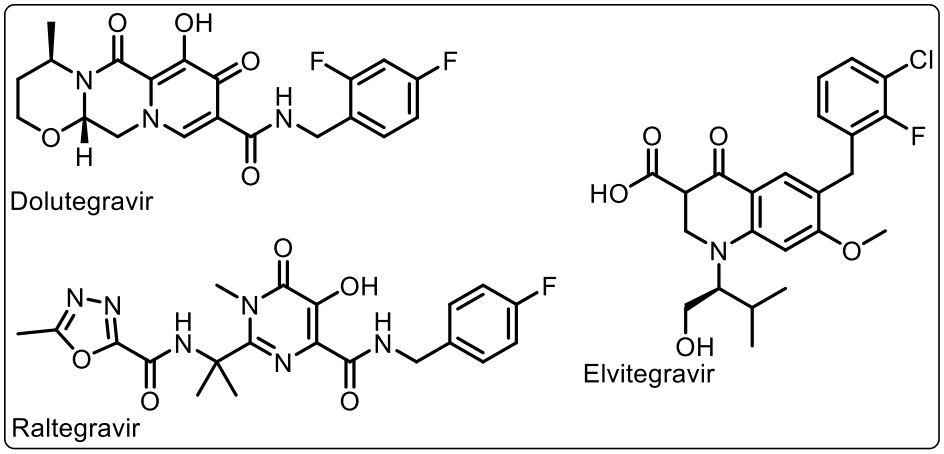


Figure 8: The integrase inhibitors dolutegravir, elvitegravir and raltegravir.

1.4.1.4 Transcription

Transcript elongation requires the binding of the HIV-1 regulatory protein Tat to the HIV-1 RNA element (TAR), a mechanism unique to HIV-1.[84] HIV transcription (Tat-TAR interaction) inhibitors have been identified, but to date none were potent and/or selective enough to pass phase I clinical trials.[36,85–87]

1.4.1.5 Virus assembly and production

Assembly and maturation of HIV-1 on the inner plasma membrane have also been targeted in drug discovery.[36] Betulinic acid, for example, blocks HIV-1 maturation through viral capsid interactions.[54,88] Insufficient antiviral activity however, disqualified this compound from further clinical trials.[89]

1.4.1.6 Protease processing

The HIV-1 protease enzyme is responsible for the cleavage of viral gag and gag-pol polyprotein precursors (peptide linkage) during the maturation of the virus.[90] HIV-1 PIs, apart from tipranavir based on the coumarin scaffold, are large peptide-like compounds with a hydroxyethylene scaffold that block proteolysis of the viral polyprotein by mimicking the normal peptide linkage.[36,52] These compounds cannot be cleaved and thus prevent the normal functioning of HIV protease ('peptidomimetic' principle).[52] They are often co-administered with "boosting" agents such as cobicistat (Tybost®) and ritonavir (RTV), which prevents the metabolism of PIs by inhibiting the activity of cytochrome P34A, thereby enhancing drug levels.[2,36,91,92] The ten PIs currently in use are: amprenavir (APV, Agenerase®), atazanavir (ATZ, Reyataz®), darunavir (DRV, Prezista®), fosamprenavir (FPV, Lexiva®), indinavir (IDV, Crixivan®), lopinavir (LPV, Kaletra®), nelfinavir (NFV, Viracept®), ritonavir (RTV, Norvir®), saquinavir (SQV, Fortovase®/ Invirase®), and tipranavir (TPV, Aptivus®) (*Figure 9*).

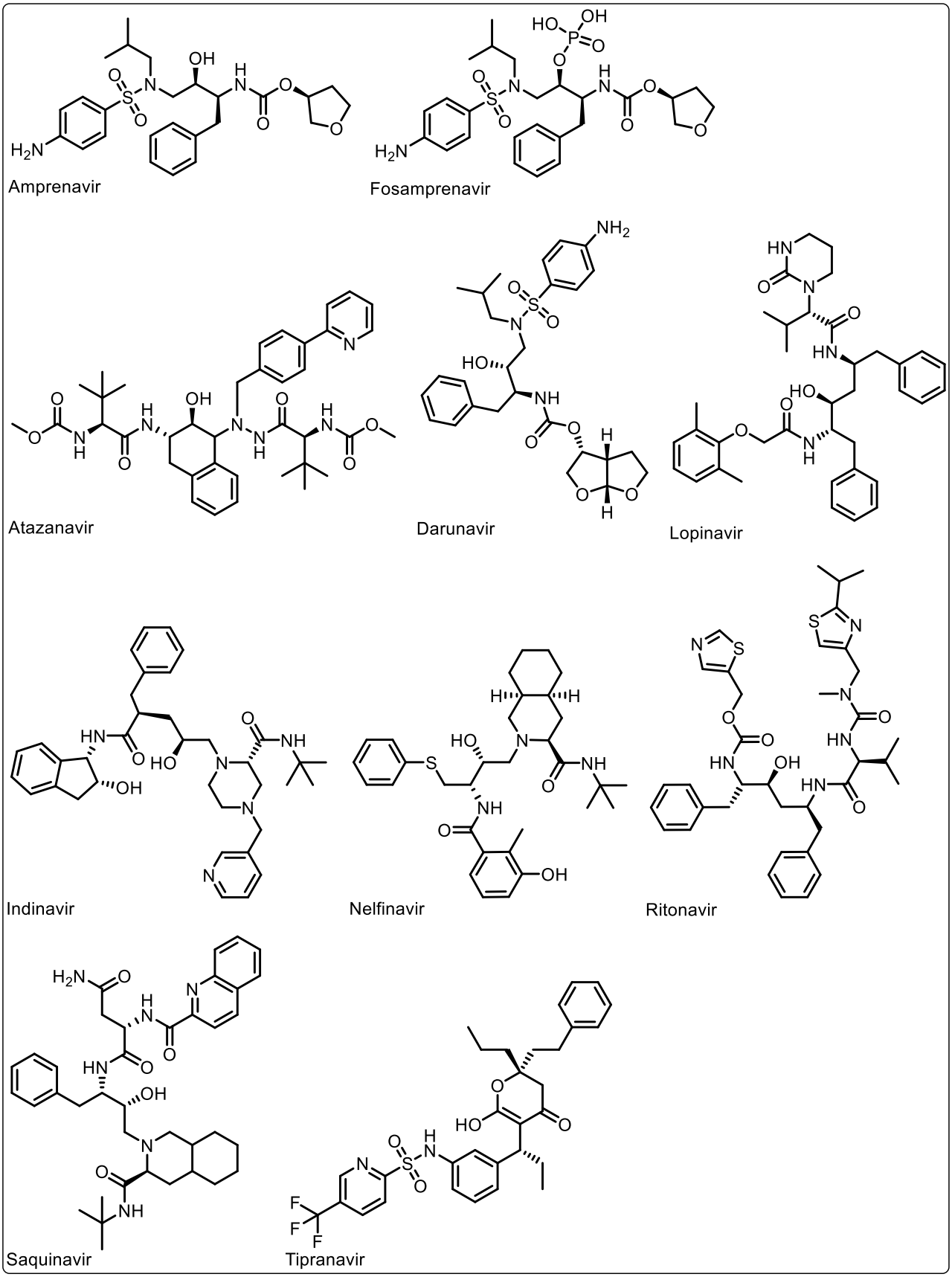


Figure 9: The many protease inhibitors used in clinical practice.

1.4.2 More on Nonnucleoside Reverse Transcriptase Inhibitors

1.4.2.1 The enzyme

Reverse transcription of the viral single-stranded (+) RNA genome into double-stranded DNA is entirely reliant on the enzymatic activities of the retroviral RT enzyme.[63,64] Crystal structural data has revealed that the RT enzyme is an asymmetric heterodimer composed of a 560-amino-acid, 66 kDa subunit (p66) and a 440-amino-acid, 51 kDa subunit (p51) (*Figure 10*). [13,63,64,93] The heterodimer has one DNA polymerization active site, that can use RNA or DNA as template, and one endonucleolytic ribonuclease H (RNase H) active site, able to degrade the RNA strand of the RNA:DNA duplex, both of which are situated in the p66 subunit at spatially distinct regions.[13,64,93] The polymerase domain resembles a hand with fingers (residues 1-85 and 118-155), palm (residues 86-117 and 156-237) and thumb (residues 238-318) subdomains. Three aspartic acid residues, D110, D185 and D186, are situated in the palm subdomain.[64,93] Additional domains include the connection (residues 319-426) and the C-terminal RNase H domain (residues 427-560).[64] The connection domain joins the polymerase and RNase H domains and is concerned with nucleic acid substrate interactions and RT inter-subunit interactions.[64,93] The RNase H active site is characterised by four conserved acidic amino acids (D443, E478, D498 and D549).[64] The p51 polypeptide subunit stemmed from HIV-1 protease-mediated cleavage of the C-terminal RNase H domain of the p66 polypeptide.[63] The same polymerase and connection domains are found in this subunit, but where p66 adopts an “open” catalytically-competent conformation that is able to accommodate a nucleic acid template strand, p51 is in a “closed” conformation rendering it non-functional.[64]

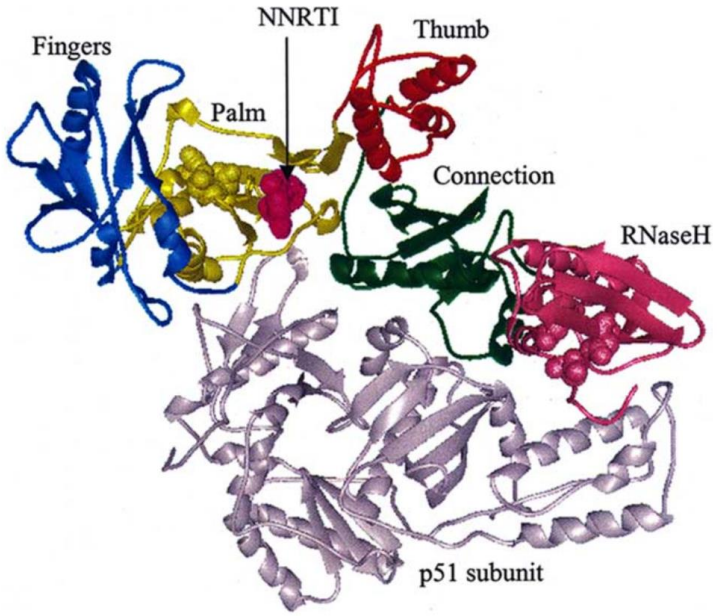


Figure 10: The structure of HIV-1 reverse transcriptase enzyme in complex with nevirapine (obtained from “Conformational Changes in HIV-1 Reverse Transcriptase Induced by Nonnucleoside Reverse Transcriptase Inhibitor Binding”). The p66 fingers (blue), palm (yellow), thumb (red) and connection (green) subdomains and the RNase H domain (pink) are illustrated.[64]

1.4.2.2 NNRTI binding pocket

NNRTIs bind to a site on the p66 subunit of the HIV-1 RT p66/p51 heterodimer called the NNRTI binding pocket (NNRTI-BP) which, interestingly, only exists in the presence of a suitable ligand.[13,63,64,94] It is situated between the β6-β10-β9 and β12-β13-β14 sheets in the palm subdomain of the p66 subunit, approximately 10 Å from the aspartic acid catalytic triad. [13,64,93] The NNRTI-BP is primarily hydrophobic with considerable aromatic character (Y181, Y188, F227, W229, and Y232). Several hydrophilic residues (K101, K103, S105, D192 and E224 of the p66 subunit and E138 of the β7-β8 loop of the p51 subunit) however also occupy this space.[63,64] Solvent may enter the binding pocket at the p66/p51 heterodimer interface, surrounded by residues L100, K101, K103, V179 and Y181 of the p66 subunit and E138 of the P51 subunit.[64] In the absence of a ligand, the side chains of Y181 and Y188 of p66 point into the hydrophobic core causing a surface depression at the accepted pocket entrance so that no pocket can be recognised. Upon ligand binding, the side chains rotate away creating a space for the ligand to occupy.[13,63,64] Once a ligand interacts with the binding pocket, short- and long-range distortions to the RT structure result. The short-range distortions include conformational changes of the amino acids and/or structural elements that form part of the pocket. Examples are the reorientation of the side chains of Y181 and Y188 and the displacement of the β12-β13-β14 sheet. The hinge-bending movement of the p66 thumb subdomain is considered a long-range distortion. This results in the displacement of the p66 connection, the RNase H domain and the p51 subunit relative to the polymerase active site.[64] NNRTIs exercise their antiviral activity by locking the active catalytic site in an inactive conformation.[73]

1.4.2.3 Ligands

The NNRTIs developed to date are small (<600 Da) hydrophobic compounds with diverse structures that, unlike NRTIs, do not require intracellular metabolism for activity.[64,94] NNRTIs are selective for HIV-1 and can be divided into two classes of compounds namely first- and second-generation compounds (*Figure 11*). First generation compounds were generally discovered by the random screening of compound libraries. These drugs are unfortunately prone to the rapid development of drug resistance.[64] They bind to amino acids Y181, Y188 and W229 on one side and K103, V106 and V179 on the other. Examples of such compounds are delavirdine, efavirenz and nevirapine (*Figure 11*, left). The development of second generation compounds, such as etravirine and rilpivirine, involves molecular modelling, rationale-based synthesis and biological and pharmacokinetic evaluations (*Figure 11*, right). Second generation compounds were developed to be more potent than first generation compounds and retain activity for longer in the case of viral mutations.[64,94] Common features of NNRTI pharmacophores important for interacting with the binding pocket are aromatic rings capable of π-stacking interactions, functional groups able to participate in hydrogen bonding, and one or more hydrocarbon-rich regions conferring hydrophobic contacts.[64]

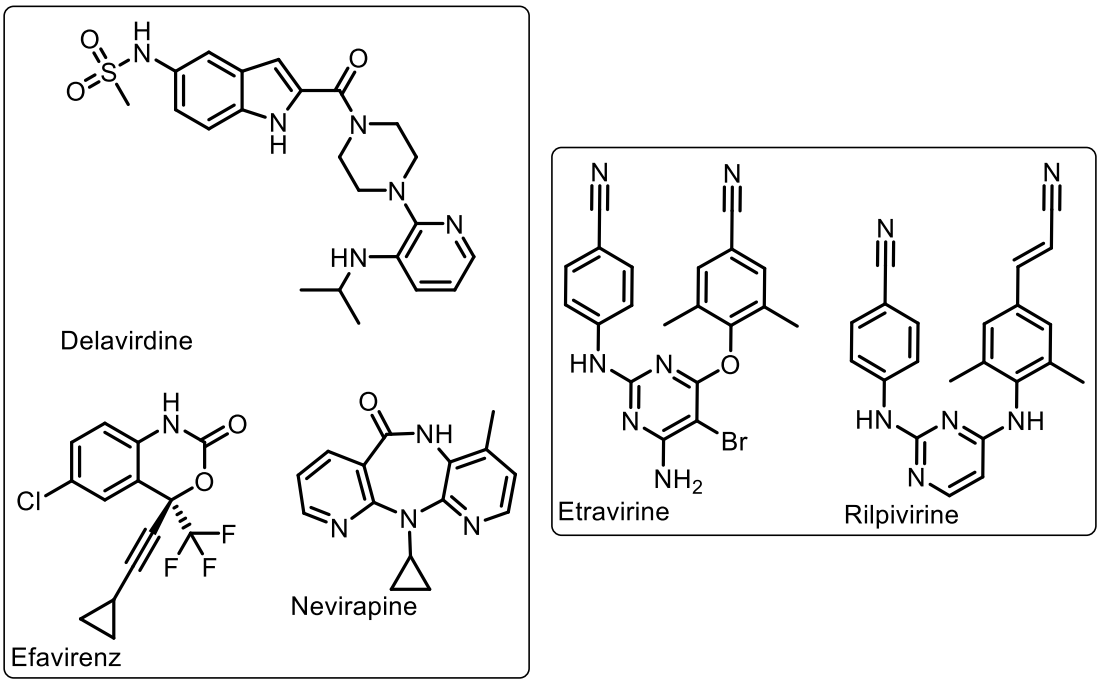


Figure 11: First (left) and second (right) generation non-nucleoside reverse transcriptase inhibitors.

1.4.2.4 HIV-1 resistance

During reverse transcription, the RT enzyme, having a weak affinity for the template, frequently "jumps" from one template to the other. If the two RNA molecules in one viron are not genetically identical, template switching will result in a novel recombinant DNA genome. The high frequency of genetic recombination, combined with the high mutation rate of HIV-1 RT (3 X 10^{-5} per cycle of replication) result in HIV populations with heterogeneous genome sequences. In this manner, the HI virus gains the ability to rapidly evade the host immune response and develop resistance to antiviral drugs.[31]

NNRTIs are particularly notorious for rapidly triggering the emergence of drug resistant HIV-1 strains, especially under monotherapy.[13,73] This causes significant problems in maintaining effective therapy. For this reason, the utility of this drug class initially seemed limited, until multi-drug combination therapy established an important role for NNRTIs. At this point in drug design, improved resistance profiles is a particular objective.[13] The risk for developing resistance is the determining major factor of the role of a NNRTI in long-term treatment.[73]

1.4.2.4.1 Mutations

NNRTI resistance develops from amino acid substitutions clustered around the binding pocket. The side-chains of amino acids undergoing mutations point towards the pocket. Resistance to NNRTIs may therefore develop via changes in protein side-chain interactions with the inhibitors. The types

of mutations that result in resistance are numerous and varied. Mutations may, for example, increase or decrease the steric bulk of the side-chains resulting in steric conflict or a loss of contact with the inhibitor. Charge may be gained or lost or the sign could be switched from positive to negative or *vice versa* which could alter the affinity of the pocket for the NNRTI.[63] Mutations mainly occur in two regions: the 'floor' of the pocket (β-sheet containing β9, β10, and β6) and the loop between β5b and β6. The 'floor' region shows mutations including V106A, V108I, V179N/E, Y181C, Y188H, and G190E, and the latter region includes mutations A98G, L100I, K101E, and K103N. The mutations L228F and P236L are located on the pocket 'roof' formed by β12, β13, and β14. Finally, E138K may be found on the loop between β7 and β8 of the p51 subunit.[63] *Figure 12* shows NNRTI resistance mutations included on the list of the International AIDS Society (www.iasusa.org).[95] Understanding the mechanism by which mutations confer resistance is important when designing new NNRTIs with the ability to maintain potency in the presence of mutations.[2] Major mutations at positions 100, 103, 181 and 188, shown in bold (*Figure 12*), occurring under the treatment of the most commonly prescribed NNRTIs, are discussed. K103N and Y181C, being the most commonly found mutations observed with virtually all NNRTIs, will be focused on.[73]

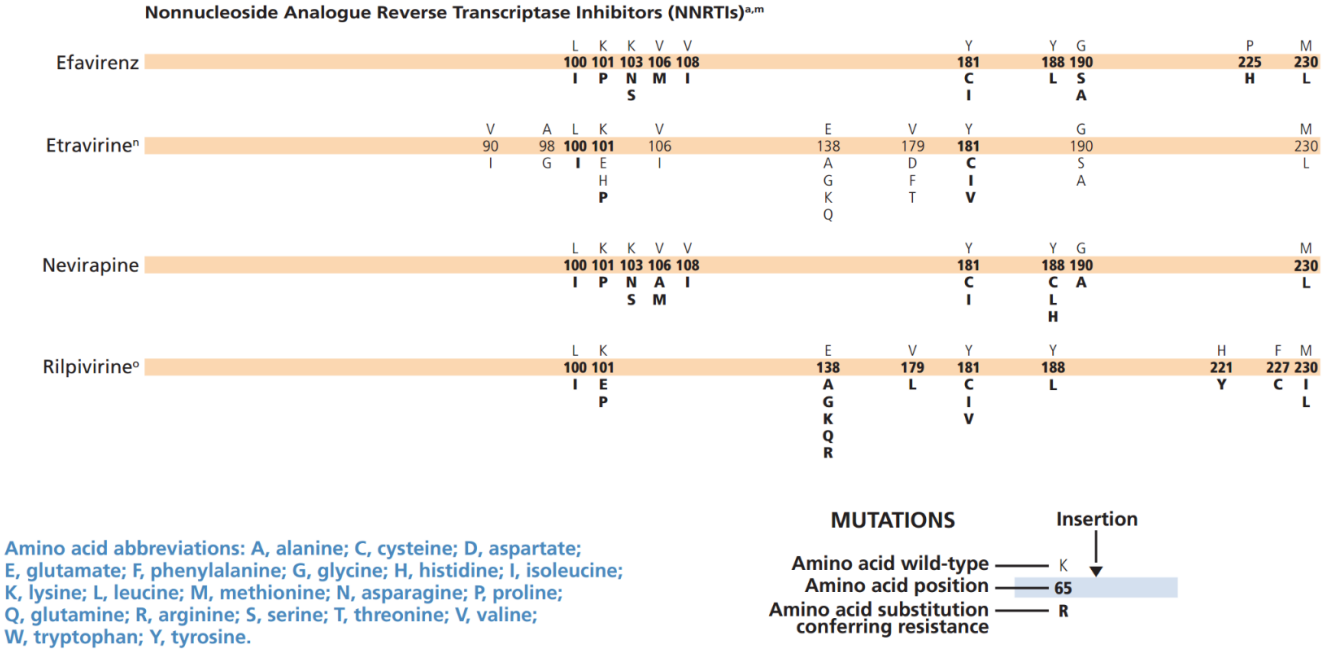


Figure 12: Mutations in the reverse transcriptase gene associated with resistance to non-nucleoside reverse transcriptase inhibitors [obtained from the International AIDS Society web page (www.iasusa.org)].[95] The most common amino acid mutations are shown in bold.

The K103N mutation is the most commonly reported mutation for NNRTIs arising from clinical use of this drug class.[13] While the mutation only minimally changes the enzyme structure, it causes a 20- to 50-fold resistance to each of the commonly used NNRTIs.[96–98] An explanation for the loss of binding potency resulted from a crystal structure of the unliganded HIV-1 RT into which K103N was modelled. It showed a hydrogen bond between the hydroxyl group of the Tyr188 side-chain in the

unliganded 'down' position and the Asn103 amide group.[13,99] This appears to stabilise the closed pocket form and interfere with the binding ability of inhibitors.[99] The NNRTI would have to overcome an increased energy barrier before binding can take place. This accounts for the reduced potency and explains why K103N encodes resistance to a very wide range of NNRTIs. The reason for the preservation of almost wild-type activity by second generation inhibitors in the presence of this mutation is still unknown. The increased stability of the unliganded conformation of the pocket may therefore only be a partial explanation for the resistance displayed.[13]

L100I usually occurs with K103N in patients receiving efavirenz and significantly increases resistance to the drug. L100I causes intermediate resistance to efavirenz and delavirdine and low-level resistance to nevirapine (*Figure 11*).[100] The conversion of leucine into isoleucine is likely to give rise to unfavourable steric interactions within the binding pocket.[2]

Significant conformational differences in the binding pocket were noted for NNRTIs with a wide range of potencies. Emivirine (MKC-442) and TNK-651 (IC_{50}s 10 nM), for example, induced the up position of Tyr181 and Tyr188 while HEPT (IC50 17 μM) had no effect on Tyr181 and did not interact optimally with Tyr188 (*Figure 13*). The conformational change can be attributed to the bulk in position 5' of the pyrimidine-2,4-dione ring. An isopropyl group in this position can reposition Tyr181 while a methyl group is not big enough to change the conformation of Tyr181 and induce favourable aromatic ring stacking.[13] Crystal structures confirmed that π-π interactions between tyrosine and the aromatic rings present in NNRTIs drive the binding of first generation NNRTIs. Mutations at these positions therefore result in a significant decrease in inhibition.[13,63] Y181C/I, where the aromatic tyrosine side chain is replaced by an aliphatic group, cause more than 30-fold resistance to nevirapine and delavirdine and 2- to 3-fold resistance to efavirenz due to the loss of aromatic ring stacking interactions (*Figure 11*).[13,96,97,101,102] Similarly, Y188L causes high level resistance to nevirapine (>1500 fold) and efavirenz (140 fold) and intermediate resistance to delavirdine (17 fold). Y188C and Y188H are uncommon mutations causing intermediate-to-high levels of nevirapine resistance and low-levels of resistance to efavirenz and delavirdine.[96–98,101,103] Second generation NNRTIs are affected to a lesser extent by these mutations since they tend to make less contact with these amino acids.[13,73]

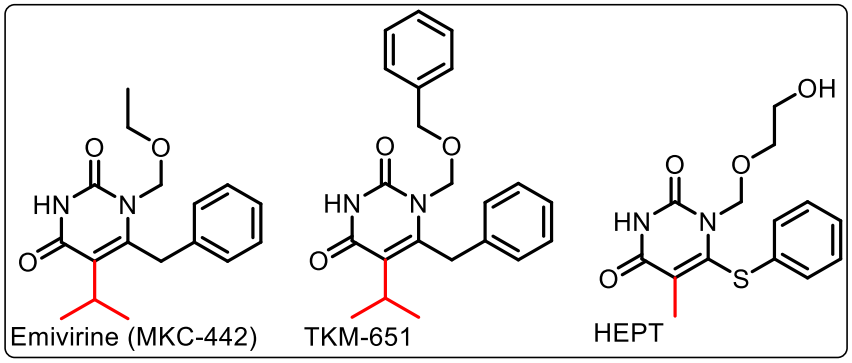


Figure 13: Non-nucleoside reverse transcriptase inhibitors inducing conformational changes to the binding pocket, with groups in position 5' of the pyrimidine-2,4-dione ring (red), resulting in different potencies.

1.4.2.4.2 Inhibitors with greater resilience to known mutations

A structure-based drug design study, aimed at creating a favourable profile of resilience to drug resistance mutations, obtained success when interactions with Tyr181 and Tyr188 with MKC-442 (*Figure 13*) were minimalised and interactions with the functionally important and highly conserved Try229 were optimised by introducing 3',5'-dimethyl substituents on the benzyl group. The new compound, GCA-186 (*Figure 14*), had improved activity against the Y181C mutant compared to the parent molecule. The same resilience is seen in efavirenz which interacts with Tyr181 and Tyr188 side chains via a small cyclopropyl group, rather than bulky pyridyl rings seen in nevirapine and delavirdine (*Figure 14*).[13] The dominant binding interaction for efavirenz is a hydrogen bonding interaction with the main chain carbonyl oxygen of Lys101. This interaction is also seen in second generation inhibitors, etravirine and rilpivirine (*Figure 11*), maintaining excellent potency against the Y181C and Y188C/L mutants.[2] Etravirine and rilpivirine additionally have flexible chemical structures which allow for several possible conformations within the binding pocket and efficient binding despite the presence of mutations.[13] In addition to its flexibility, capravirine (*Figure 14*) form extensive main-chain H-bonds involving the residues K101, K103 and P236 of the p66 subunit.[13,73] The inhibitor is less likely to be influenced by mutations compared to when side-chain interactions are utilised.[13]

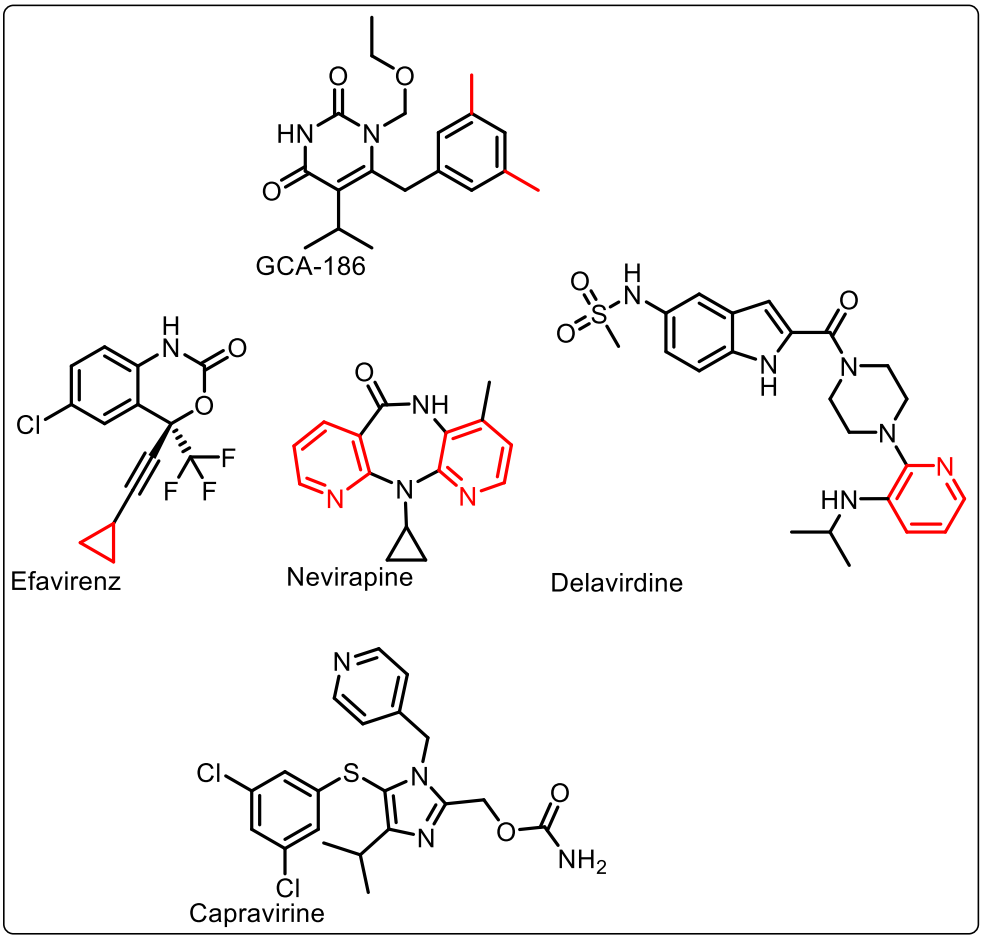


Figure 14: Structural features resulting in resilience to drug resistant mutations. Important groups are highlighted in red.

1.4.2.5 Paving the path forward

A list of 25 different FDA approved HIV drugs, as well as 14 drugs formulated as combinations, are currently available for the treatment of HIV infection.[104] The development of effective drug-resistance testing has enabled the tailoring of treatments for patients with varying resistance profiles.[45] Unfortunately, no antiretroviral drug combination is completely effective in shutting down viral replication. The virus thus continues to diversify and challenge the antiretroviral drug regimen used.[45] The remaining problems regarding the treatment of HIV are therefore drug resistance, transmissibility and side effects of the current treatment regimens. The development of new treatments, because of the need for life-long therapy, is clearly critical.[11] Maintained activity against mutations, ease of synthesis, oral viability and prolonged duration of activity should be aimed for when designing and synthesising new drugs. Further needs to be met are bringing therapy and better health infrastructure to poor nations, continuing and advancing education, global monitoring of the different strains of HIV for changes in their virulence and development of a preventive vaccine.[52]

With the role of NNRTIs in HAART firmly established together with abundant information on the RT enzyme and the NNRTI-BP, this drug class makes for an excellent springboard towards a new era of superior drugs against HIV.

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less. ~ Marie Curie

1.5 Bibliography

1 R. C. Gallo and L. Montagnier, *Science*, 2002, **298**, 1730–1731.

2 C. Reynolds, C. B. de Koning, S. C. Pelly, W. A. L. van Otterlo and M. L. Bode, *Chem. Soc. Rev.*, 2012, **41**, 4657–4670.

3 UNAIDS, *UNAIDS fact sheet: Latest statistics on the status of the AIDS epidemic*, 2017, www.unaids.org., Accessed 16/08/2017.

4 Unaids, *Unaids South Africa*, 2016,www.unaids.org., Accessed 18/08/2017.

5 J. Thurlow, J. Gow and G. George, *J. Int. AIDS Soc.*, 2009, **12**, 1–13.

6 A. Boutayeb, *BMC Public Health*, 2009, **9**, 1–10.

7 A. Pouris and A. Pouris, *Scientometrics*, 2011, **86**, 541–552.

8 R. C. Gallo, *Science*, 2002, **298**, 1728–1730.

9 R. C. Gallo and L. Montagnier, *N. Engl. J. Med.*, 2003, **349**, 2283–2285.

10 R. C. Gallo, *Transfusion*, 2015, **55**, 1–9.

11 R. C. Gallo, *Retrovirology*, 2006, **3**, 1–7.

12 S. Martins, M. J. Ramos and P. A. Fernandes, *Curr. Med. Chem.*, 2008, **15**, 1083–1095.

13 J. Ren and D. K. Stammers, *Virus Res.*, 2008, **134**, 157–170.

14 P. M. Sharp and B. H. Hahn, *Nature*, 2008, **455**, 605–606.

15 N. Vidal, M. Peeters, C. Mulanga-Kabeya, N. Nzilambi, D. Robertson, W. Ilunga, H. Sema, K. Tshimanga, B. Bongo and E. Delaporte, *J. Virol.*, 2000, **74**, 10498–10507.

16 P. M. Sharp and B. H. Hahn, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 2010, **365**, 2487–2494.

17 V. M. Hirsch, R. A. Olmsted, M. Murphey-Corb, R. H. Purcell and P. R. Johnson, *Nature*, 1989, **339**, 389–392.

18 N. L. Letvin, M. D. Daniel, P. K. Sehgal, R. C. Desrosiers, R. D. Hunt, L. M. Waldron, J. J. MacKey, D. K. Schmidt, L. V. Chalifoux and N. W. King, *Science*, 1985, **230**, 71–73.

19 C. Apetrei, A. Kaur, N. W. Lerche, M. Metzger, I. Pandrea, J. Hardcastle, S. Falkenstein, R. Bohm, J. Koehler, V. Traina-Dorge, T. Williams, S. Staprans, G. Plauche, R. S. Veazey, H. McClure, A. A. Lackner, B. Gormus, D. L. Robertson and M. A. Preston, *J. Virol.*, 2005, **79**, 8991–9005.

20 F. Gao, L. Yue, A. T. White, P. G. Pappas, J. Barchue, A. P. Hanson, B. M. Greene, P. M. Sharp, G. M. Shaw and B. H. Hahn, *Nature*, 1992, **358**, 495–499.

21 B. H. Hahn, G. M. Shaw, K. M. De Cock and P. M. Sharp, *Science*, 2000, **287**, 607–614.

22 M. Peeters, C. Honore, T. Huet, L. Bedjabaga, S. Ossari, P. Bussi, R. W. Cooper and E. Delaporte, *AIDS*, 1989, **3**, 625–630.

23 T. Huet, R. Cheynier, A. Meyerhans, G. Roelants and S. Wain-Hobson, *Nature*, 1990, **345**, 356–359.

24 M. Peeters, K. Fransen, E. Delaporte, M. Van den Haesevelde, G.-M. Gershy-Damet, L. Kestens, G. van der Groen and P. Piot, *AIDS*, 1992, **6**, 447–451.

25 F. Gao, E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp and B. H. Hahn, *Nature*, 1999, **397**, 436–441.

26 A. M. Prince, B. Brotman, D.-H. Lee, L. Andrus, J. Valinsky and P. Marx, *AIDS Res. Hum. Retroviruses*, 2002, **18**, 657–660.

27 W. M. Switzer, B. Parekh, V. Shanmugam, V. Bhullar, S. Phillips, J. J. Ely and W. Heneine, *AIDS Res. Hum. Retroviruses*, 2005, **21**, 335–342.

28 B. F. Keele, J. H. Jones, K. A. Terio, J. D. Estes, R. S. Rudicell, M. L. Wilson, Y. Li, G. H. Learn, T. M. Beasley, J. Schumacher-Stankey, E. Wroblewski, A. Mosser, J. Raphael, S. Kamenya, E. V. Lonsdorf, D. A. Travis, T. Mlengeya, M. J. Kinsel, J. G. Else, G. Silvestri, J. Goodall, P. M. Sharp, G. M. Shaw, A. E. Pusey and B. H. Hahn, *Nature*, 2009, **460**, 515–519.

29 Centers for Disease Control and Prevention, 2017, www.cdc.gov/hiv/transmission.html., Accessed 21/08/2017.

30 E. M. Campbell and T. J. Hope, *Trends Microbiol.*, 2008, **16**, 580–587.

31 E. O. Freed, *Somat. Cell Mol. Genet.*, 2001, **26**, 13–33.

32 N. Sluis-Cremer, D. Arion and M. A. Parniak, *Cell. Mol. Life Sci.*, 2000, **57**, 1408–1422.

33 T. D. Hollingsworth, R. M. Anderson and C. Fraser, *J. Infect. Dis.*, 2008, **198**, 687–693.

34 V. Dahl, L. Josefsson and S. Palmer, *Antiviral Res.*, 2010, **85**, 286–294.

35 S. Jaffar, A. D. Grant, J. Whitworth, P. G. Smith and H. Whittle, *Bull. World Health Organ.*, 2004, **82**, 462–469.

36 E. J. Arts and D. J. Hazuda, *Cold Spring Harb. Perspect. Med.*, 2012, **2**, 1–24.

37 J. P. Horwitz, J. Chua and M. Noel, *J. Org. Chem.*, 1964, **29**, 2076–2078.

38 H. Mitsuya, K. J. Weinhold, P. A. Furman, M. H. St. Clair, S. Nusinoff Lehrman, R. C. Gallo, D. Bolognesi, D. W. Barry and S. Broder, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 7096–7100.

39 B. Autran, G. Carcelain, T. S. Li, C. Blanc, D. Mathez, R. Tubiana, C. Katlama, P. Debre and J. Leibowitch, *Science*, 1997, **277**, 112–116.

40 K. V. Komanduri, M. N. Viswanathan, E. D. Wieder, D. K. Schmidt, B. M. Bredt, M. A.

Jacobson and J. M. McCune, *Nat. Med.*, 1998, **4**, 953–956.

41 M. M. Lederman, E. Connick, A. Landay, D. R. Kuritzkes, J. Spritzler, M. St. Clair, B. L. Kotzin, L. Fox, M. Heath Chiozzi, J. M. Leonard, F. Rousseau, M. Wade, J. D'Arc Roe, A. Martinez and H. Kessler, *J. Infect. Dis.*, 1998, **178**, 70–79.

42 L. M. Mansky and H. M. Temin, *J. Virol.*, 1995, **69**, 5087–5094.

43 P. K. O'Neil, G. Sun, H. Yu, Y. Ron, J. P. Dougherty and B. D. Preston, *J. Biol. Chem.*, 2002, **277**, 38053–38061.

44 M. E. Abram, A. L. Ferris, W. Shao, W. G. Alvord and S. H. Hughes, *J. Virol.*, 2010, **84**, 9864–9878.

45 F. Clavel, *Physicians Res. Netw. Noteb.*, 2004, **9**, 3–7.

46 J. M. Coffin, *Science*, 1995, **267**, 483–489.

47 S. D. W. Frost and R. McLean, Angela, *AIDS*, 1994, **8**, 323–332.

48 M. A. Nowak, S. Bonhoeffer, G. M. Shaw and R. M. May, *J. Theor. Biol.*, 1997, **184**, 203–217.

49 R. F. Stengel, *Math. Biosci.*, 2008, **213**, 93–102.

50 F. Maggiolo, *J. Antimicrob. Chemother.*, 2009, **64**, 910–928.

51 AIDS info, 2016, http://aidsinfo.nih.gov., Accessed 23/08/2017.

52 E. De Clercq, *Int. J. Antimicrob. Agents*, 2009, **33**, 307–320.

53 K. A. Reimann, W. Lin, S. Bixler, B. Browning, B. N. Ehrenfels, J. Lucci, K. Miatkowski, D. Olson, T. H. Parish, M. D. Rosa, F. B. Oleson, Y. Ming Hsu, E. A. Padlan, N. L. Letvin and L. C. Burkly, *AIDS Res. Hum. Retroviruses*, 1997, **13**, 933–943.

54 P.-F. Lin, W. Blair, T. Wang, T. Spicer, Q. Guo, N. Zhou, Y.-F. Gong, H.-G. H. Wang, R. Rose, G. Yamanaka, B. Robinson, C.-B. Li, R. Fridell, C. Deminie, G. Demers, Z. Yang, L. Zadjura, N. Meanwell and R. Colonno, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 11013–11018.

55 D. R. Kuritzkes, J. Jacobson, W. G. Powderly, E. Godofsky, E. DeJesus, F. Haas, K. A. Reimann, J. L. Larson, P. O. Yarbough, V. Curt and W. R. Shanahan, *J. Infect. Dis.*, 2004, **189**, 286–291.

56 T. Dragic, A. Trkola, D. a. D. Thompson, E. G. Cormier, F. A. Kajumo, E. Maxwell, S. W. Lin, W. Ying, S. O. Smith, T. P. Sakmar and J. P. Moore, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 5639–5644.

57 F. Tsamis, S. Gavrilov, F. Kajumo, C. Seibert, S. Kuhmann, T. Ketas, A. Trkola, A. Palani, J. W. Clader, J. R. Tagat, S. McCombie, B. Baroudy, J. P. Moore, T. P. Sakmar and T. Dragic, *J. Virol.*, 2003, **77**, 5201–5208.

58 R. S. Veazey, P. J. Klasse, S. M. Schader, Q. Hu, T. J. Ketas, M. Lu, P. A. Marx, J. Dufour, R. J. Colonno, R. J. Shattock, M. S. Springer and J. P. Moore, *Nature*, 2005, **438**, 99–102.

59 R. M. Gulick, J. Lalezari, J. Goodrich, N. Clumeck, E. DeJesus, A. Horban, J. Nadler, B. Clotet, A. Karlsson, M. Wohlfeiler, J. B. Montana, M. McHale, J. Sullivan, C. Ridgway, S. Felstead, M. W. Dunne, E. van der Ryst and H. Mayer, *N. Engl. J. Med.*, 2017, **359**, 1429–1441.

60 C. B. Wilen, J. C. Tilton and R. W. Doms, *Cold Spring Harb. Perspect. Med.*, 2012, **2**, 1–13.

61 J. A. Esté and A. Telenti, *Lancet*, 2007, **370**, 81–88.

62 T. Matthews, M. Salgo, M. Greenberg, J. Chung, R. DeMasi and D. Bolognesi, *Nat. Rev. Drug Discov.*, 2004, **3**, 215–225.

63 C. Tantillo, J. Ding, A. Jacobo-Molina, R. G. Nanni, P. L. Boyer, S. H. Hughes, R. Pauwels, K. Andries, P. A. J. Janssen and E. Arnold, *J. Mol. Biol.*, 1994, **243**, 369–387.

64 N. Sluis-cremer, N. A. Temiz and I. Bahar, *Curr. HIV Res.*, 2004, **2**, 323–332.

65 P. A. Furman, J. A. Fyfe, M. H. St. Clair, K. Weinhold, J. L. Rideout, G. A. Freeman, S. Nusinoff Lehrman, D. P. Bolognesi, S. Broder, H. Mitsuya and D. W. Barry, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 8333–8337.

66 H. Mitsuya and S. Broder, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 1911–1915.

67 M. H. St. Clair, C. A. Richards, T. Spector, K. J. Weinhold, W. H. Miller, A. J. Langlois and P. A. Furman, *Antimicrob. Agents Chemother.*, 1987, **31**, 1972–1977.

68 G. J. Hart, D. C. Orr, C. R. Penn, H. T. Figueiredo, N. M. Gray, R. E. Boehme and J. M. Cameron, *Antimicrob. Agents Chemother.*, 1992, **36**, 1688–1694.

69 A. S. Ray, M. W. Fordyce and M. J. M. Hitchcock, *Antiviral Res.*, 2016, **125**, 63–70.

70 E. De Clercq, *Biochem. Pharmacol.*, 2016, **119**, 1–7.

71 C. Wyatt and J. M. Baeten, *Lancet*, 2015, **385**, 2559–2560.

72 L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice and T. A. Steitz, *Science*, 1992, **256**, 1783–1790.

73 E. De Clercq, *Chem. Biodivers.*, 2004, **1**, 44–64.

74 R. Esnouf, J. Ren, C. Ross, Y. Jones, D. Stammers and D. Stuart, *Struct. Biol.*, 1995, **2**, 303–308.

75 M. Baba, H. Tanaka, E. De Clercq, R. Pauwels, J. Balzarini, D. Schols, H. Nakashima, C.-F. Perno, R. T. Walker and T. Miyasaka, *Biochem. Biophys. Res. Commun.*, 1989, **165**, 1375–1381.

76 R. Pauwels, K. Andries, J. Desmyter, D. Schols, M. J. Kukla, H. J. Breslin, A. Raeymaeckers, J. Van Gelder, R. Woestenborghs, J. Heykants, K. Schellekens, M. A. C. Janssen, E. De Clercq and P. A. J. Janssen, *Nature*, 1990, **343**, 470–474.

77 A. S. Espeseth, P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J. Y. Melamed, S. Young, T. Hamill, J. L. Cole and D. J. Hazuda, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11244–11249.

78 D. J. Hazuda, N. J. Anthony, R. P. Gomez, S. M. Jolly, J. S. Wai, L. Zhuang, T. E. Fisher, M. Embrey, J. P. Guare, M. S. Egbertson, J. P. Vacca, J. R. Huff, P. J. Felock, M. V. Witmer, K. A. Stillmock, R. Danovich, J. Grobler, M. D. Miller, A. S. Espeseth, L. Jin, I.-W. Chen, J. H. Lin, K. Kassahun, J. D. Ellis, B. K. Wong, W. Xu, P. G. Pearson, W. A. Schleif, R. Cortese, E. Emini, V. Summa, M. K. Holloway and S. D. Young, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**,

11233–11238.

79 D. J. Hazuda, S. D. Young, J. P. Guare, N. J. Anthony, R. P. Gomez, J. S. Wai, J. P. Vacca, L. Handt, S. L. Motzel, H. J. Klein, G. Dornadula, R. M. Danovich, M. V. Witmer, K. A. A. Wilson, L. Tussey, W. A. Schleif, L. S. Gabryelski, L. Jin, M. D. Miller, D. R. Casimiro, E. A. Emini and J. W. Shiver, *Science*, 2004, **305**, 528–532.

80 D. J. McColl and X. Chen, *Antiviral Res.*, 2010, **85**, 101–118.

81 J. A. Grobler, K. Stillmock, B. Hu, M. Witmer, P. Felock, A. S. Espeseth, A. Wolfe, M. Egbertson, M. Bourgeois, J. Melamed, J. S. Wai, S. Young, J. Vacca and D. J. Hazuda, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 6661–6666.

82 M. Sato, T. Motomura, H. Aramaki, T. Matsuda, M. Yamashita, Y. Ito, H. Kawakami, Y. Matsuzaki, W. Watanabe, K. Yamataka, S. Ikeda, E. Kodama, M. Matsuoka and H. Shinkai, *J. Med. Chem.*, 2006, **49**, 1506–1508.

83 K. Shimura, E. Kodama, Y. Sakagami, Y. Matsuzaki, W. Watanabe, K. Yamataka, Y. Watanabe, Y. Ohata, S. Doi, M. Sato, M. Kano, S. Ikeda and M. Matsuoka, *J. Virol.*, 2008, **82**, 764–774.

84 J. Karn and M. C. Stoltzfus, *Cold Spring Harb. Perspect. Med.*, 2012, **4**, 116–133.

85 M. Hsu, A. D. Schutr, M. Holly, L. W. Slice, M. I. Sherman, D. D. Richman, M. Jane Potash and D. J. Volsky, *Science*, 1991, **254**, 1799–1802.

86 F. Hamy, E. R. Felder, G. Heizmann, J. Lazdins, F. Aboul-Ela, G. Varani, J. Karn and T. Klimkait, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 3548–3553.

87 S. Hwang, N. Tamilarasu, K. Kibler, H. Cao, A. Ali, Y.-H. Ping, K.-T. Jeang and T. M. Rana, *J. Biol. Chem.*, 2003, **278**, 39092–39103.

88 T. Fujioka and Y. Kashiwada, *J. Nat. Prod.*, 1994, **57**, 243–247.

89 P. F. Smith, A. Ogundele, A. Forrest, J. Wilton, K. Salzwedel, J. Doto, G. P. Allaway and D. E. Martin, *Antimicrob. Agents Chemother.*, 2007, **51**, 3574–3581.

90 J. Park and C. D. Morrow, *Virology*, 1993, **194**, 843–850.

91 J. Kempf, Dale, K. C. Marsh, G. Kumar, D. A. Rodrigues, F. Denissen, Jon, E. McDonald, M. J. Kukulka, A. Hsu, R. G. Granneman, P. A. Baroldi, E. Sun, D. Pizzuti, J. J. Plattner, D. W. Norbeck and J. M. Leonard, *Antimicrob. Agents Chemother.*, 1997, **41**, 654–660.

92 A. Hsu, G. R. Granneman, G. Cao, L. Carothers, T. El-Shourbagy, P. Baroldi, K. Erdman, F. Brown, E. Sun and J. M. Leonard, *Clin. Pharmacol. Ther.*, 1998, **63**, 453–464.

93 L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice and T. A. Steitz, *Science*, 1992, **256**, 1783–1790.

94 J. Ghosn, M.-L. Chaix and C. Delaugerre, *AIDS Rev.*, 2009, **11**, 165–173.

95 International Antiviral Society - USA, HIV Drug Resistance Mutations Figures and User Notes, 20117, http://www.iasusa.org/content/drug-resistance-mutations-in-HIV, Accessed 26/08/2017.

96 S. D. Young, S. F. Britcher, L. O. Tran, L. S. Payne, W. C. Lumma, T. A. Lyle, J. R. Huff, P. S. Anderson, D. B. Olsen, S. S. Carroll, D. J. Pettibone, J. A. O'Brien, R. G. Ball, S. K. Balani, J. H. Lin, I. Chen, W. A. Schleif, V. V. Sardana, W. J. Long, V. W. Byrnes and E. A. Emini, *Antimicrob. Agents Chemother.*, 1995, **39**, 2602–2605.

97 C. J. Petropoulos, N. T. Parkin, K. L. Limoli, Y. S. Lie, T. Wrin, W. Huang, H. Tian, D. Smith, G. A. Winslow, D. J. Capon and J. M. Whitcomb, *Antimicrob Agents Chemother.*, 2000, **44**, 920–928.

98 L. Bacheler, S. Jeffrey, G. Hanna, R. D'Aquila, L. Wallace, K. Logue, B. Cordova, K. Hertogs, B. Larder, R. Buckery, D. Baker, K. Gallagher, H. Scarnati, R. Tritch and C. Rizzo, *J. Virol.*, 2001, **75**, 4999–5008.

99 Y. Hsiou, J. Ding, K. Das, A. D. Clark Jr, P. L. Boyer, P. Lewi, P. A. J. Janssen, J.-P. Kleim, M. Rösner, S. H. Hughes and E. Arnold, *J. Mol. Biol.*, 2001, **309**, 437–445.

100 L. T. Bacheler, E. D. Anton, P. Kudish, D. Baker, J. Bunville, K. Krakowski, L. Bolling, M. Aujay, X. Victoria Wang, D. Ellis, M. F. Becker, A. L. Lasut, H. J. George, D. R. Spalding, G. Hollis and K. Abremski, *Antimicrob Agents Chemother.*, 2000, **44**, 2475–2484.

101 V. W. Byrnes, V. V. Sardana, W. A. Schleif, J. H. Condra, J. A. Waterbury, J. A. Wolfgang, W. J. Long, C. L. Schneider, A. J. Schlabach, B. S. Wolanski, D. J. Graham, L. Gotlib, A. Rhodes, D. L. Titus, E. Roth, O. M. Blahy, J. C. Quintero, S. Staszewski and E. A. Emini, *Antimicrob. Agents Chemother.*, 1993, **37**, 1576–1579.

102 V. W. Byrnes, E. A. Emini, W. A. Schleif, J. H. Condra, C. L. Schneider, W. J. Long, J. A. Wolfgang, D. J. Graham, L. Gotlib, A. J. Schlabach, B. S. Wolanski, O. M. Blahy, J. C. Quintero, A. Rhodes, E. Roth, D. L. Titus and V. V. Sardana, *Antimicrob Agents Chemother.*, 1994, **38**, 1404–1407.

103 T. Fujiwara, A. Sato, M. El-Farrash, S. Miki, K. Abe, Y. Isaka, M. Kodama, Y. Wu, L. B. Chen, H. Harada, H. Sugimoto, M. Hatanaka and Y. Hinuma, *Antimicrob. Agents Chemother.*, 1998, **42**, 1340–1345.

104 AIDS info, FDA-Approved HIV Medicines, 2017, https://aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/21/58/fda-approved-hiv-medicines, Accessed 28/08/2017.

Chapter 2: Indole-derived Non-nucleoside Reverse Transcriptase Inhibitors

2.1 Our strategy

In this section, the work that served as inspiration for this portion of the project is briefly discussed and the target compounds for the synthetic part of the research project, are introduced.

2.1.1 Interesting indoles

The NNRTI-BP presents opportunities for creative probing, with the structures of known NNRTIs varying significantly.[1] The NNRTI, efavirenz (*Figure 2.1*), was recognised as a particularly interesting starting point for the design of new, more potent drug candidates based on the double hydrogen bond interaction between the amide portion of the carbamate and Lys101 in the binding pocket. The formation of these same interactions was anticipated with an indole-2-carboxamide based structure and a literature search for this scaffold, applied as an NNRTI, was conducted.[1] This lead to the discovery that in 1993, a phenylsulfinyl-indole (**1**, *Figure 2.1*) active against HIV-1 RT was reported.[2] Pursuing this success, many groups investigated the potential of this class of compound as drug candidate. Thus far, several variations of the first indole-based NNRTI have been developed. These include motifs that contain sulfonyls (**2**, *Figure 2.1*), sulphonamides (**3**) and phosphinates (**4**).[3–7]

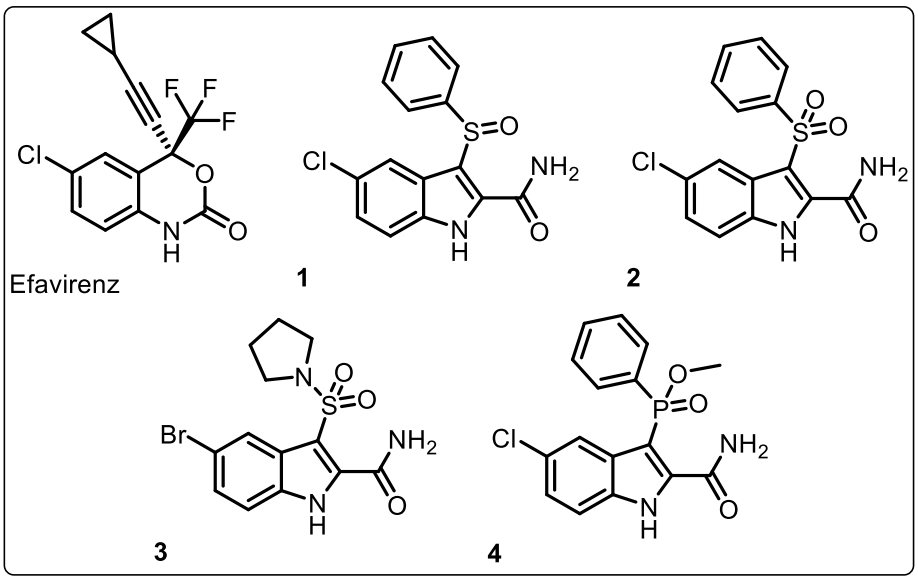


Figure 2.1: Efavirenz and examples of the indole scaffold as NNRTIs.

The pharmacologist and Nobel laureate, James Black once advised that “the most fruitful basis for the discovery of a new drug is to start with an old drug”.[8] In light of the successes of indole derivatives as potent NNRTIs, it was proposed that this scaffold could be optimised by the introduction of new functional groups in the 2-, 3- and 5-positions of the indole, while retaining the hydrogen bonding

interactions to Lys101 through the indole NH and a suitable hydrogen bond acceptor at the indole 2-position.[1]

2.1.1.1 Important interactions

Already discussed are the two hydrogen bonding interactions to Lys101 (*Figure 2.2*). Additional favourable interactions of the indole-2-carboxamide scaffold include π stacking interactions between the phenyl ring in the indole 3-position and Tyr181, as well as a σ-π interaction between the para-phenyl proton and the indole side chain of Trp229.[1] The Val179 pocket was identified as an area with potential for improvement of compound activity since the original sulfonyl group shows no electrostatic interactions to receptor residues and, in our opinion, is not particularly suitable given the hydrophobic nature of the pocket.[1]

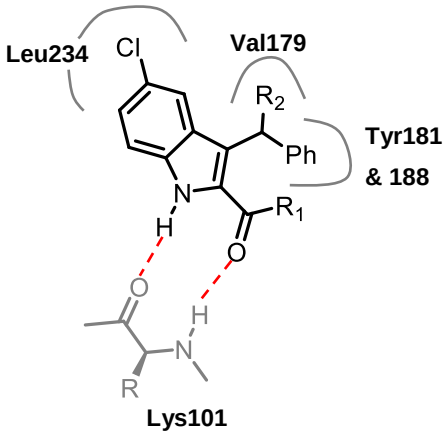


Figure 2.2: A schematic representation of the indole scaffold in the NNRTI allosteric site.

2.1.2 The building blocks

The first indole-based NNRTI synthesised in our laboratories was the carboxylate **5** (*Figure 2.3*) with a cyclopropyl group well accommodated in the Val179 pocket (as per modelling). This proof of concept molecule proved to be a poor inhibitor with an IC_{50} value of 28.5 μM, likely due to poor membrane permeability given the negatively charged carboxylate group at physiological pH. Satisfyingly, the ester analogue **6** (*Figure 2.3*), with an IC_{50} value of 0.085 μM, was an equally effective inhibitor as nevirapine.[1] Molecular modelling also showed that the hydrophobic ethyl chain protrudes out of the binding pocket into the solvation region. It was therefore envisaged that a polar group, able to interact with water molecules in the solvation region, might be more suitable at this position. Compounds **7**-**9** (*Figure 2.3*) were, however, significantly less potent than **6** (*Table 2.1*). It was concluded that the interactions with water in fact resulted in a decreased binding affinity due to attractive forces between the molecule and the solvent.[9]

The installation of the cyclopropyl group was a modestly yielding reaction, difficult to reproduce and generally rather problematic. It was also limiting since installation of this group was ineffective without the ester group in the 2-position. A Structure Activity Relationship (SAR) study was then carried out

to determine the importance of occupying the Val179 pocket and indeed, it had an effect on the potency (*Figure 2.3*, *Table 2.1*, **10**). One other variation involved the removal of the phenyl group (*Figure 2.3*, **11**) which also resulted in a significant deterioration of potency owing to the absence of π-π stacking interactions with Tyr181 (*Table 2.1*). Based on these results, the decision was made to retain the important phenyl group and incorporate a more easily attainable bioisostere in place of the cyclopropyl moiety. After extensive molecular modelling studies, both the *R* and *S* enantiomers of the methyl-ether analogue **12** (*Figure 2.3*), appeared promising and resulted in an increase in potency of nearly two orders of magnitude compared to **6** (*Table 2.1*).[9]

Improving the potency against common mutant strains was the next objective. A known strategy is the incorporation of methyl substituents on the 3- and 5-positions of a phenyl ring occupying the Tyr181/188 pocket. To this end, compound **13** (*Figure 2.3*) was synthesised and both **12** and **13** were tested against a range of mutant strains. The dimethyl analogue performed better across all mutant strains even though there was no significant difference in potency between the two compounds against the wild type HIV-1 (*Table 2.1*). Of particular interest was the potency of **13** against the most prevalent mutation, K103N. While compound **12** was rendered ineffective against this mutant, **13** displayed only a slight decrease in potency.[9]

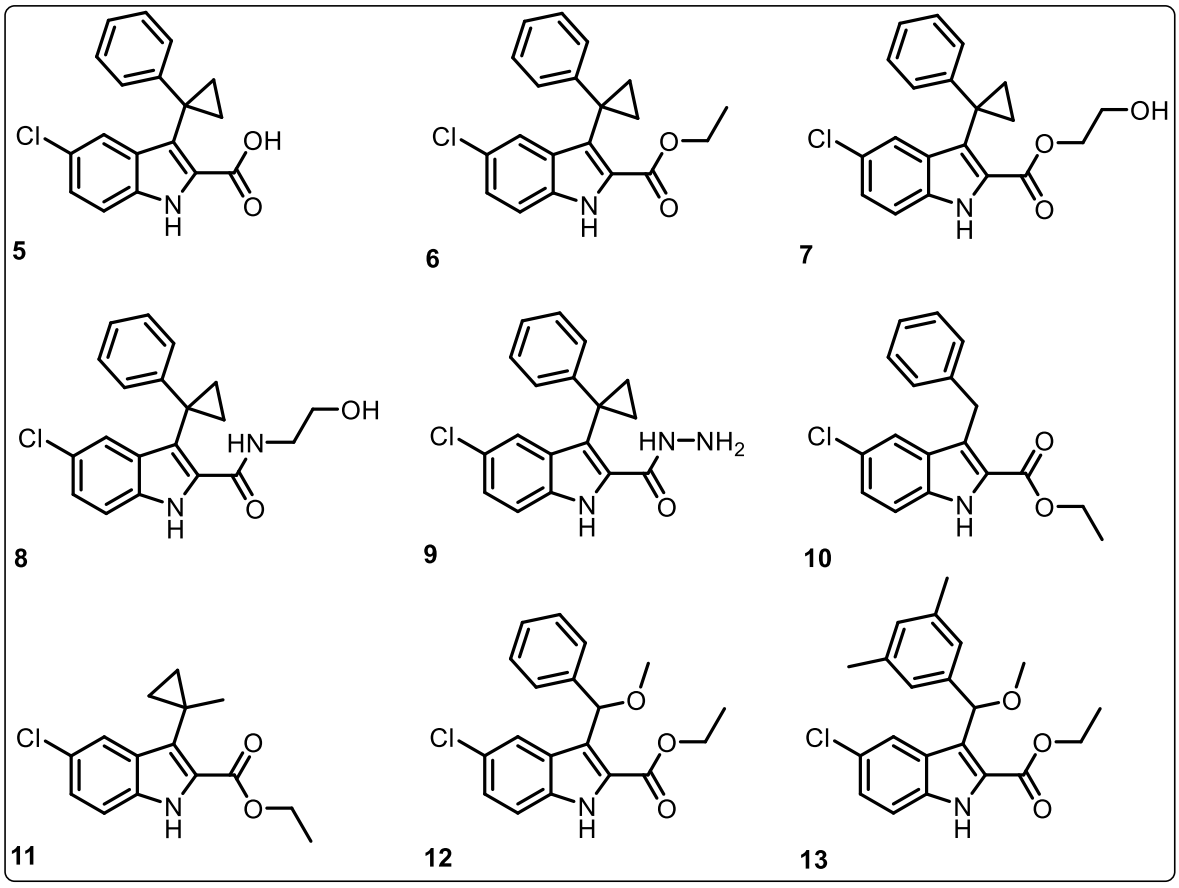


Figure 2.3: Indole based NNRTIs synthesised in our laboratories.

Table 2.1:The phenotypic assay values of indole derivatives against wild type HIV-1, the corresponding cellular toxicity values and a measure of the compound's hydrophilicity (compounds shown in Figure 2.3).[1,9] *IC_{50}: The concentration of the inhibitor when binding is reduced by half. CC_{50}: The concentration of a compound at which cell viability is reduced by half. cLogP: A measurement of a compound's hydrophilicity. A high logP value corresponds to low hydrophilicity and thus poor absorption or permeation. This value should ideally not exceed 5.0.*

Compound	IC_{50}/μM	CC_{50}/μM	cLogP
5	28.5	92.3	-
6	0.085	30.3	4.8
7	0.712	-	3.9
8	0.209	-	3.4
9	4.40	-	2.8
10	0.24	>100	-
11	2.25	44.3	-
12	0.02	25.3	-
13	0.03	36.5	-

While the indole moiety showed great potential as a scaffold for NNRTIs, a significant problem was noticed with the methyl ether derivatives **12** and **13**. Under acidic conditions, such as those in the stomach, the methyl ether could be activated and indole nitrogen-facilitated elimination of methanol would occur. Subsequent attack by water could produce the corresponding alcohol, a poor inhibitor of HIV-1 RT (*Figure 2.4*). This class of compounds would therefore be expected to lose efficacy after oral administration.[10]

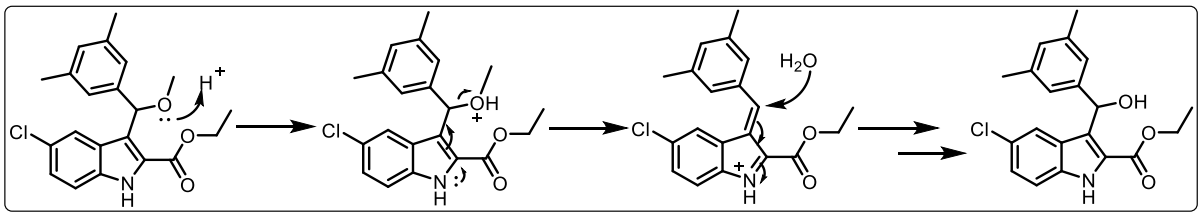


Figure 2.4:The possible acid-mediated degradation of methyl ether drug candidates by an indole nitrogen-assisted S_N1 mechanism.

Methyl sulfide bioisosteres (see compounds **14** and **15**, *Figure 2.5*) were then synthesised as replacements of the methyl ether analogues. These proved to be slightly poorer inhibitors than their methyl ether counterparts (*Table 2.2*), but were found to be much less susceptible to acid promoted degradation, with less than 5% conversion of the sulphide into the ethyl ether in ethanol and concentrated sulfuric acid after 72 hours. These observations made sense since sulphide is a poorer Lewis base and thus more stable compounds could be expected.[10]

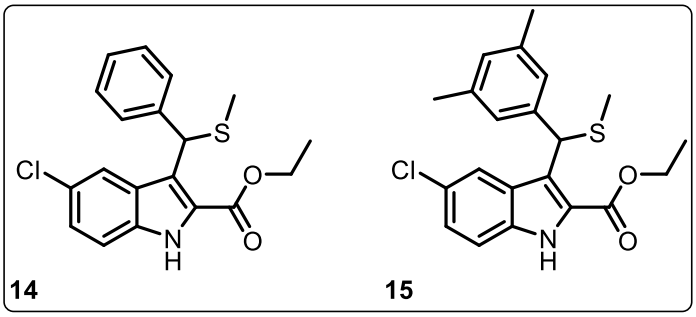


Figure 2.5: Methyl sulfide bioisosteres of the methyl ether indole based NNRTIs.

Table 2.2: The phenotypic assay values of the methyl sulfide bioisosteres against wild type HIV-1 and their corresponding cellular toxicity values.[10]

Compound	IC_{50}/μM	CC_{50}/μM
14	0.039	26.1
15	0.060	32.0

2.1.2.1 A natural next step

With the methyl sulfide being larger than the methyl ether, this group may not be accommodated as well by the small Val179 pocket, resulting in a decrease in potency. Furthermore, the electronegative ether group draws electron density from the indole scaffold, making the indole NH a better hydrogen bond donor than in the case of a methyl sulfide group.[10] Although the sulfide derivatives proved to be much more stable under acidic conditions, degradation was not eliminated completely.[10] Replacing this moiety with an alkyl chain of appropriate length, for accommodation in the pocket, may be a more sensible strategy. With carbon being smaller than both oxygen and sulfur, this group may also fit more comfortably in the pocket. The synthesis of a small library of such compounds, probing the Val179 pocket, was planned. Since most indole NNRTIs in literature were synthesised with an amide in the 2-position (see Section 2.1.1), amide and ester derivatives were designed for comparative purposes. *Figure 2.6* shows a set of compounds (**16** – **27**) to be synthesised.

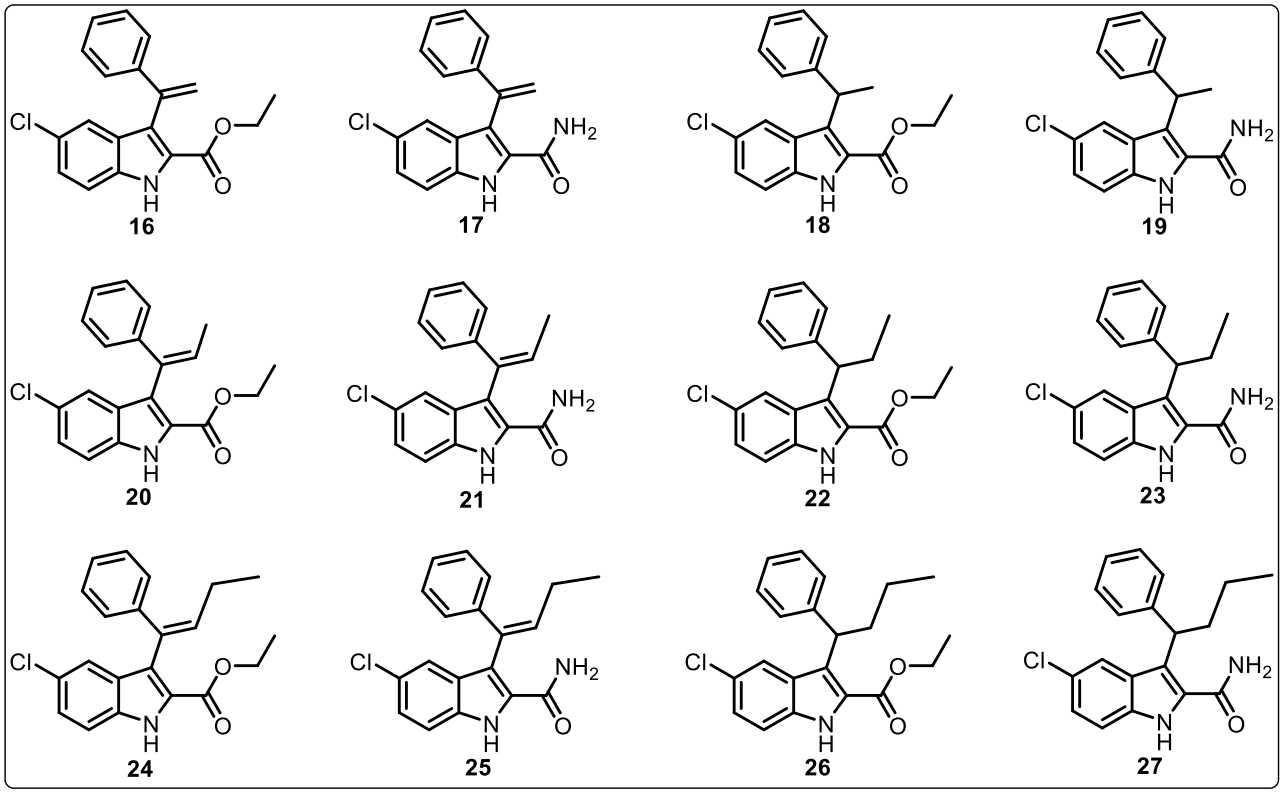


Figure 2.6: A target library of compounds with the purpose of probing the Val179 pocket and for comparing ester and amide 2-position derivatives.

2.1.3 Exploring the solvent accessible cleft

Silvestri and co-workers have focused their efforts toward the development of indolylarylsulfone (IAS) NNRTIs and have identified three regions of the IAS scaffold important for the development of potent inhibitors against both wild type and mutant HIV-1 strains (*Figure 2.7*). These are: a) the introduction of two methyl groups at positions 3 and 5 of the 3-phenylsulfonyl moiety, increased the potency against resistant strains (*Figure 2.7*, **A**); b) a chlorine atom at position 5 of the indole ring or a fluorine at position 4 enhanced activity (*Figure 2.7*, **B**); and c) potent inhibitors resulted when either a natural or an unnatural amino acid was coupled to the indole-2-carboxamide (*Figure 2.7*, **C**).[11,12]

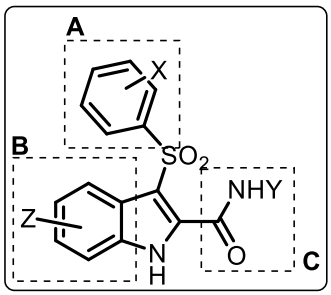


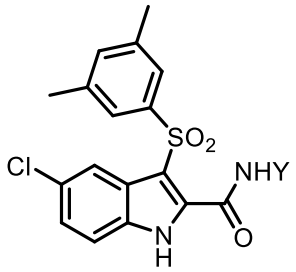
Figure 2.7: Three important regions in the IAS scaffold.

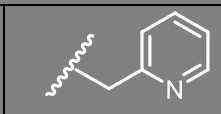
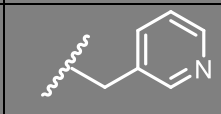
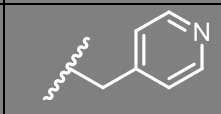
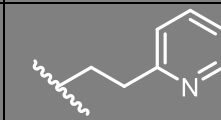
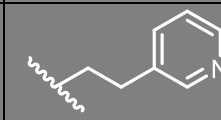
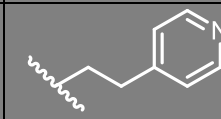
Since the first indole NNRTI, various groups have been introduced at the indole 2-position (*Figure 2.7*, **C**), not all being equally successful.[2–6,11,12] A need for strategic implementation of groups at this position arose. Silvestri and co-workers therefore performed numerous docking studies inside the NNRTI-BP of RT.[12] IASs bearing heterocyclic substituents at the indole-2-carboxamide moiety

adopted novel binding interactions in the solvent accessible cleft.[12] This cleft is composed of Arg172, Ile180, Val179, Glu138 and Thr139. It was envisaged that the introduction of a basic nitrogen atom in the side chain would establish interactions with Glu138 and that hydrophobic groups would form nonpolar interactions with the hydrophobic side chains of residues in the cleft.[12] Having recognised a novel binding interaction, this group focused their attention on exploring the **C** region further (*Figure 2.7*).[12–16] Particularly promising results were obtained when a pyridinyl moiety, linked through an methylene or ethylene bridge, was introduced as **Y** (*Table 2.3*).[13] Low nanomolar range concentrations were obtained, independent of the position of the piperidinyl nitrogen atom and all compounds performed better than the known drugs nevirapine (NVP) and efavirenz (EFV). As anticipated, the indole NH formed a hydrogen bonding interaction with Lys101, the chlorine atom pointed into a hydrophobic pocket surrounded by Val106 and Leu234, the 3,5-dimethylphenyl moiety was situated in a hydrophobic pocket formed by Tyr181, Tyr188, Trp229 and Pro95, and the pyridinyl ring formed hydrophobic interactions with the side chains of Val179 (p66), Glu138 and Thr139 (p51).

The most potent compounds, **29**, **30** and **32**, were also tested against mutant HIV-1 strains carrying the Y181C, Y188L and K103N point mutations. They were found to be potent inhibitors of the mutant Y181C strain and were superior to EFV with EC_{50}s of 8.8 nM, 2.2 nM and 6.3 nM respectively. Similarly, **29** and **30** were 7- and 14-fold more potent than EFV against the Y188L mutant strain. Importantly, these inhibitors performed well against the K103N mutation, being more effective than EFV. Compound **30** was identified as the inhibitor holding the most promise being 71-, 14-, and 8-fold more potent than EFV against Y181C, Y188L and K103N mutant strains respectively.

Table 2.3: The phenotypic assay values of IAS derivatives against wild type HIV-1, the corresponding cellular toxicity values and the selectivity index calculated as CC_{50}/EC_{50} ratio (data obtained from "New Nitrogen Containing Substituents at the Indole-2-carboxamide Yield High Potent and Broad Spectrum Indolylarylsulfone HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitors").[13]



	Y	EC_{50} (nM)	CC_{50} (nM)	SI
28		2.0 ± 0.06	>11015	>5508
29		2.2 ± 0.20	>11015	>5007
30		2.0 ± 0.20	>11015	>5508
31		2.0 ± 0.40	>10684	>5342
32		2.1 ± 0.70	>10684	>5088
33		2.0 ± 1.00	>10684	>5342
NVP		1502 ± 939	>18776	>13
EFV		6.3 ± 3.2	>15839	>2514

2.1.3.1 A continuation on previous successes

Building on positive results achieved in our laboratories and the work done by Silvestri and co-workers, incorporation of a pyridinyl moiety at the 2-position of the methyl sulfide indole derivative, was deemed a good idea (*Figure 2.8*). A small library of picolylamine and aminoethyl pyridine derivatives was designed and a synthetic route was planned. Due to the cost of the amines, specifically 4-(2-aminoethyl)pyridine (**39**, R 4130 per 1 g), the synthesis of five out of the desired six compounds was planned.

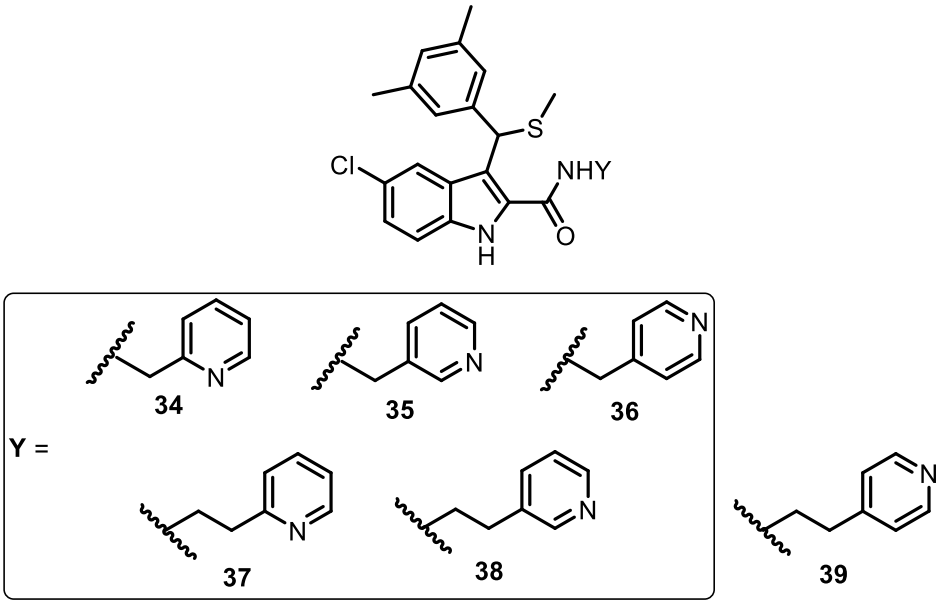


Figure 2.8: A target library of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1H-indole-2-carboxamide derivatives bearing a pyridinyl moiety linked through an methylene or ethylene bridge.

2.2 Acid stable substituted indoles

Here, the work introduced in Section 2.1.2.1 is reported. Firstly, a strategy for the synthesis of the target compounds is presented. Each reaction, followed by the synthesis, problems encountered and the successes accomplished are then discussed. Finally, the biological results are presented. It should be noted that an experimental section can be found at the end of the chapter.

2.2.1 Planning a synthetic route

The required phenyl functionality in compounds **5** to **9** (*Figure 2.9*, compounds discussed in Section 2.1.2) was readily installed by means of a Friedel-Crafts acylation reaction by Hassam *et al.* and Müller, starting from commercially available ethyl 5-chloroindole-2-carboxylate and benzoyl chloride. Next, the conversion of the ketone into an alkene was achieved through the use of a triphenylphosphonium bromide derivative and a suitable base in a Wittig reaction.[1,17] With this part of the synthesis well established, the retrosynthesis, depicted in *Scheme 2.1*, could be achieved through the more detailed synthetic plan shown in *Scheme 2.2*. Here, the Friedel-Crafts acylation reaction made use of the Lewis acid, aluminium trichloride ($AlCl_3$), while the suggested base for ylide formation in the Wittig reaction was *n*-butyllithium (*n*-BuLi).

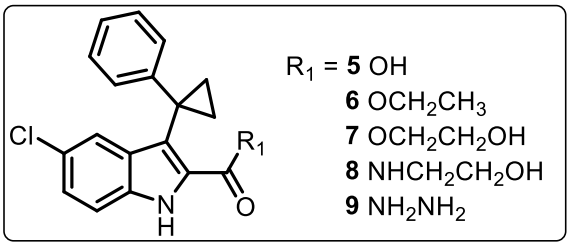
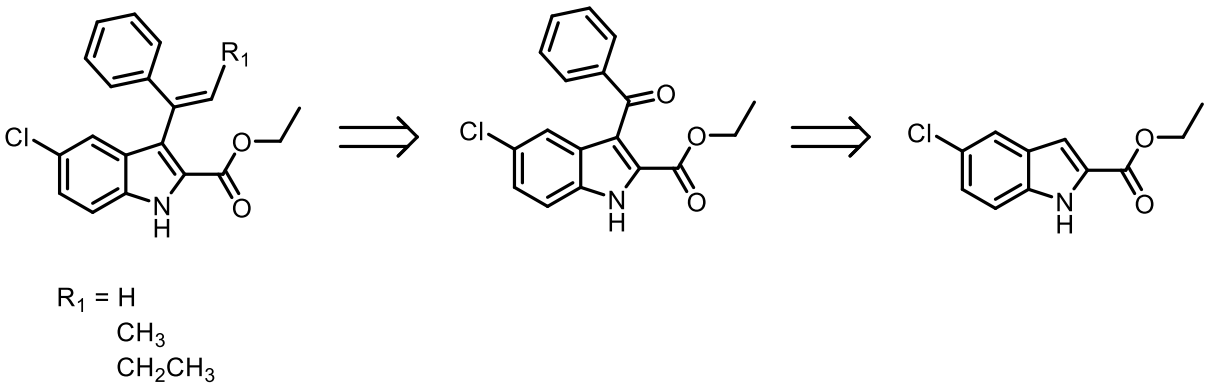
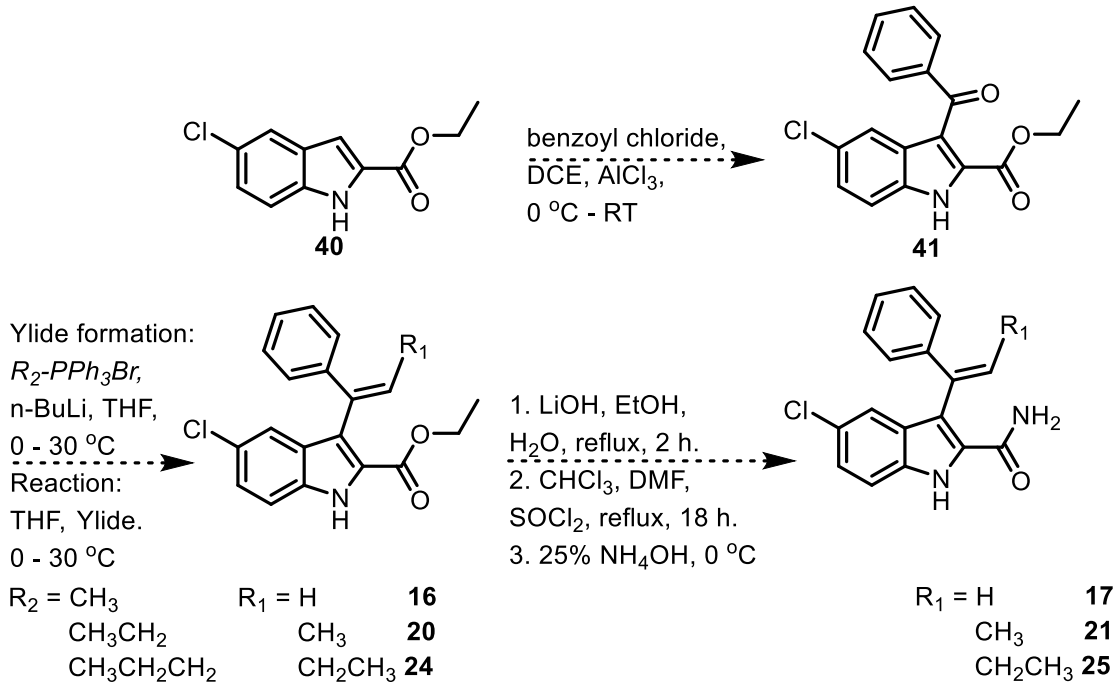


*Figure 2.9: Compounds **5** to **9**, discussed in Section 2.1.2, containing a phenyl functionality.*



Scheme 2.1: Retrosynthetic approach toward molecules designed to probe the Val179 binding pocket.

Next, the easy conversion of the ester, in the indole 2-position (**16**, **20**, **24**, *Scheme 2.2*), into the corresponding carboxylic acid was anticipated by means of base hydrolysis with lithium hydroxide (LiOH). The carboxylic acid could then be converted into the corresponding amide by conversion into the acid chloride, using thionyl chloride ($SOCl_2$), followed by quenching with aqueous ammonia (25% NH_4OH) (**17**, **21**, **25**, *Scheme 2.2*).[18]

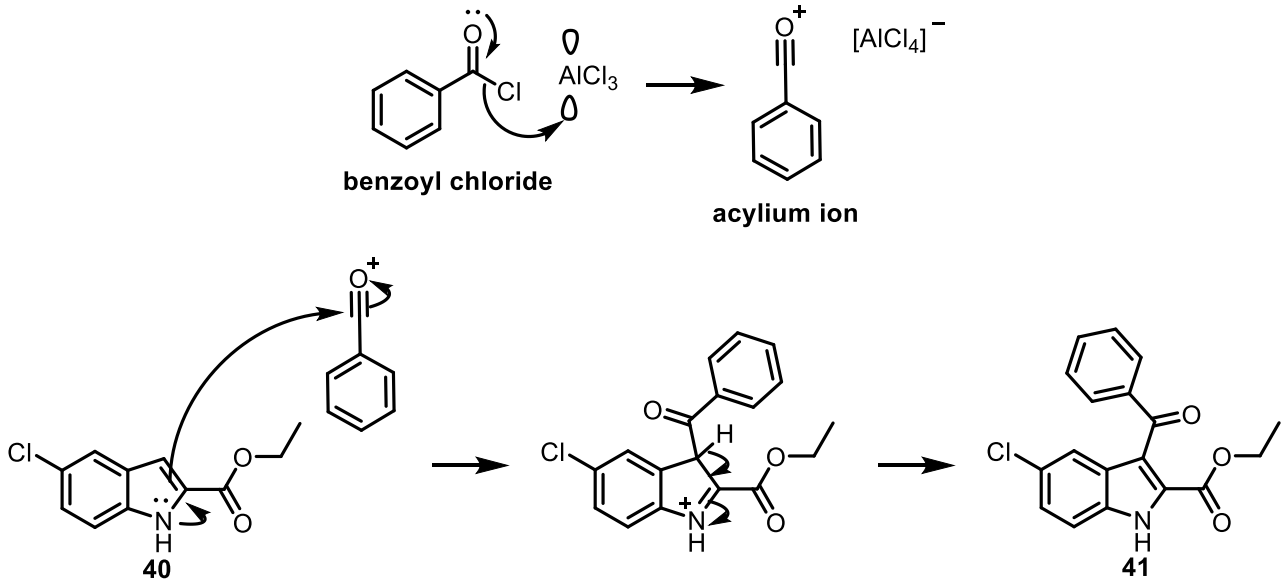


Scheme 2.2: A synthesis plan.

2.2.2 The Friedel-Crafts acylation reaction

One of the first descriptions of a Lewis acid used in organic synthesis happens to have coincided with the first example of a Friedel-Crafts alkylation reaction in 1887, in which amyl-benzene resulted from the treatment of amyl chloride with aluminium trichloride in benzene.[19] This reaction was subsequently named after its inventors, Charles Friedel and James Mason Crafts and became one of the most effective strategies for Ar – C carbon-carbon bond formation.[19,20] Unsurprisingly, this is also a common method employed to functionalise the 3-position of an indole.[21]

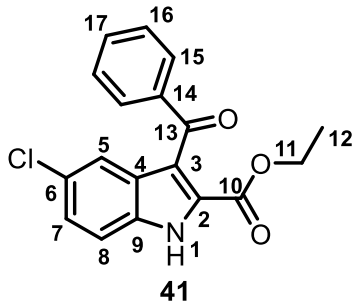
The reaction mechanism of any Friedel-Crafts reaction starts with the formation of an active species, widely accepted to be the acylium ion (*Scheme 2.3*).[20] This ion forms upon halogen transfer between an acid chloride and a Lewis acid. Over the years, many Lewis acids, among which $AlCl_3$, BF_3, $BeCl_2$, $TiCl_4$, $SbCl_5$ and $SnCl_4$, have been used for this reaction.[22] The lone pair-mediated attack by the indole on the acylium ion, causes a disruption of the indole aromaticity, resulting in a cationic intermediate. Aromaticity is regained after abstraction of a proton from the acyl-bearing sp^3 carbon of the indole, resulting in the desired reaction product **41** (*Scheme 2.3*).



Scheme 2.3: The mechanism of the Friedel-Crafts acylation reaction utilised to functionalise the indole 3-position.

2.2.2.1 Synthesis of ethyl 3-benzoyl-5-chloro-1*H*-indole-2-carboxylate (41)[1,9,10,17]

The first step in the synthetic route, a Friedel-Crafts acylation reaction, was carried out without difficulty. Anhydrous aluminium chloride ($AlCl_3$) was added to a flask containing dry 1,2-dichloroethane (DCE), followed by the dropwise addition of benzoyl chloride. Since the formation of the acylium ion is a highly exothermic reaction, low temperatures were maintained with an ice-water bath. The reaction mixture stirred for 30 minutes, to allow for the formation of the acylium ion, before the commercially available indole, ethyl 5-chloro-1*H*-indole-2-carboxylate (**40**), was added. Heating at reflux was then performed for 4 hours. During this time, the reaction mixture changed colour from a yellow to a dirty green. In the work-up, an aqueous sodium bicarbonate solution (aq. $NaHCO_3$) was added and the resulting emulsion was filtered through Celite before extraction with ethyl acetate (EtOAc) could be done. The product was obtained in good yield (72%) after purification by column chromatography.



41

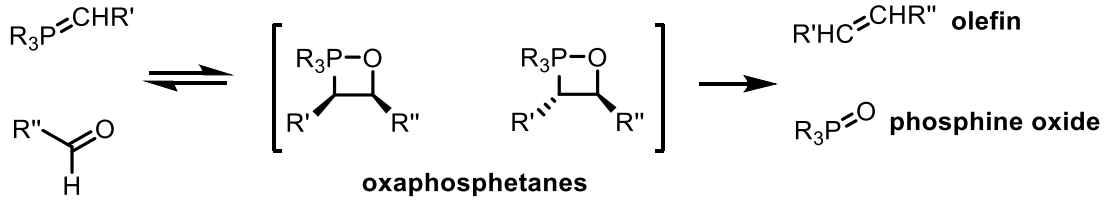
The ^{1}H nuclear magnetic resonance (NMR) spectrum, as well as the ^{13}C NMR spectrum, corresponded well to those reported in the literature.[1,9,10] A carbon signal important for indicating product formation is that of C=O observed at 192.5 ppm in the ^{13}C NMR spectrum of the product. This data, together with the observation of all protons and carbons in the benzoyl moiety, led to the conclusion that the benzoyl moiety had been successfully added at the 3-position of the indole.

2.2.3 The Wittig reaction

Around the middle of the previous century, George Wittig described a reaction in which benzophenone could be converted into diphenylethylene using methylenetriphenylphosphorane.[23,24] Since then, the Wittig reaction has been used as one of the standard methods to assemble molecules by C=C bond formation.[24] This reaction is especially attractive due to its simplicity, convenience, efficiency and stereocontrol.[25]

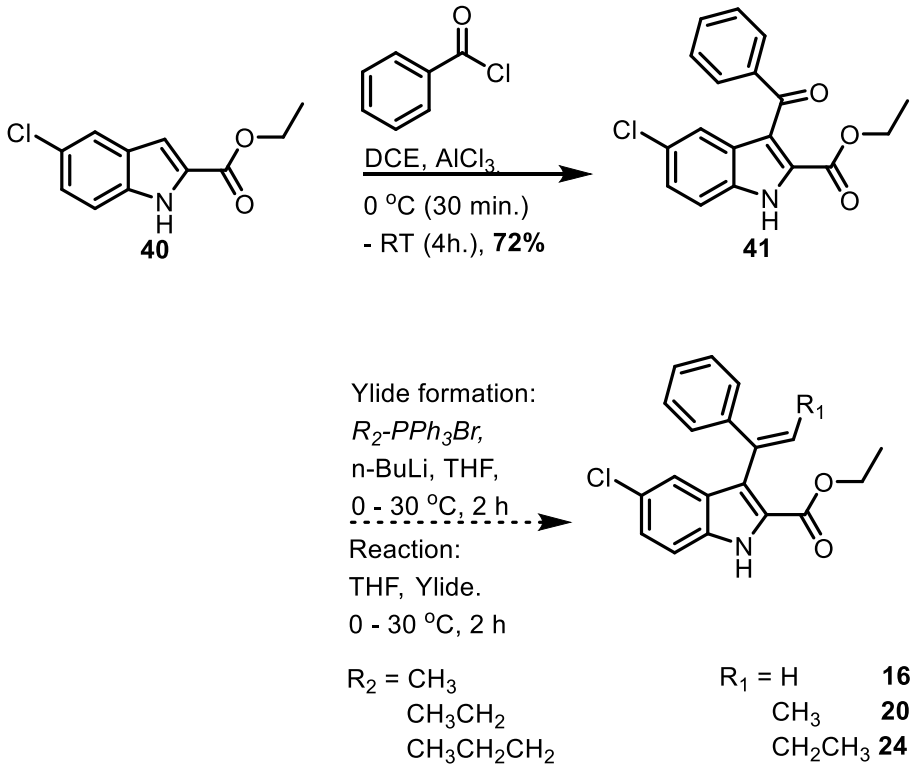
Selectivity for (*Z*)- or (*E*)-alkene products depends on the type of ylide, the type of carbonyl compound and the reaction conditions. In particular, "stabilised" ylides, having strongly conjugated substituents such as COOMe, CN or SO_2PH on the ylidic carbon, favour the formation of (*E*)-alkenes, while "non-stabilised" ylides favour the (*Z*)-alkene. "Semi-stabilised" ylides, with mildly conjugated substituents such as Ph or allyl, may form either the (*E*)- or (*Z*)-alkene.[25]

The mechanism of this reaction involves the condensation between a phosphorus ylide and an aldehyde or ketone to produce an olefin and a phosphine oxide (*Scheme 2.4*).[26] Vedejs reported for the first time in 1973 that oxaphosphetanes (*Scheme 2.4*) are the sole observable intermediates by ^{31}P NMR spectroscopy in this reaction.[27]



Scheme 2.4: The mechanism of the Wittig olefination reaction.

The Wittig reaction was used as a key step in the synthetic route, introducing variety in the indole 3-position for the purpose of probing the Val179 pocket. Through the use of methyl-, ethyl- or propyltriphenylphosphonium bromide (R_2–PPh_3Br, *Scheme 2.5*), a suitable base and the 3-benzoyl indole **41** (*Scheme 2.3*), derivatives **16**, **20** and **24** were anticipated (*Scheme 2.5* and *Figure 2.10*).



*Scheme 2.5: A proposed synthetic method for the generation of compounds **16**, **20** and **24** (shown in Figure 2.10) from 3-benzoyl indole **41**.*

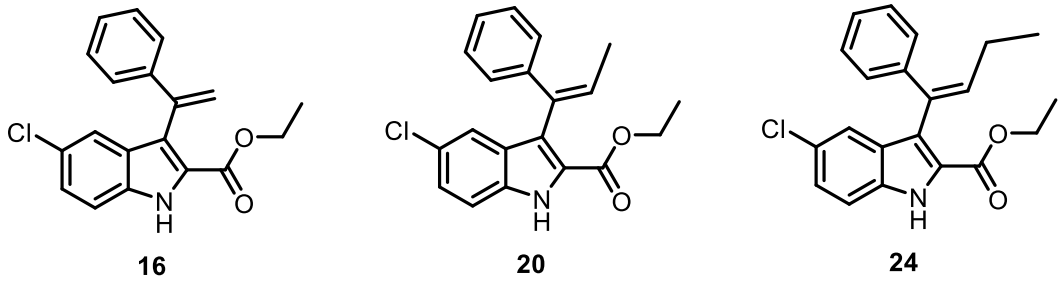


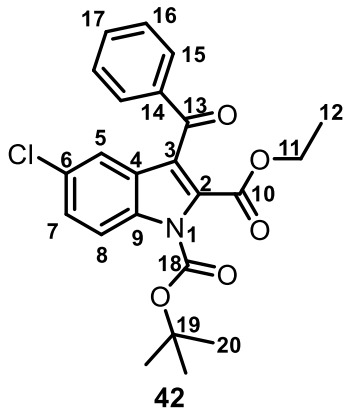
Figure 2.10: Anticipated Wittig reaction products.

2.2.3.1 Synthesis of ethyl 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (16), ethyl 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxylate (20) and ethyl 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylate (24)[1,9,17]

Dry tetrahydrofuran (THF) was cooled to 0 °C in preparation for rapid ylide formation which was subsequently achieved through the slow addition of *n*-Butyllithium (*n*-BuLi) to the THF after methyltriphenylphosphonium bromide had been added. The reaction mixture was then slowly heated to 30 °C and was stirred at this temperature for 2 hours to warrant a complete reaction. During this time, the reaction mixture turned dark red. In a separate three-neck round-bottom flask fitted with a dropping funnel, the indole starting material, ethyl 3-benzoyl-5-chloro-1*H*-indole-2-carboxylate (**41**), was dissolved in dry THF. This solution too was cooled to 0 °C, this time in preparation for exothermic cleavage of the oxaphosphatane. To this end, the solution containing the newly formed ylide was added dropwise to the indole-containing solution. After addition, the reaction mixture was stirred at 30 °C for two hours. Thin layer chromatography (TLC) showed one spot with an R_f value

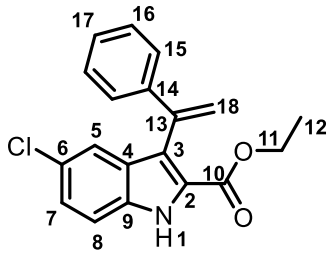
corresponding to that of the starting material. Nevertheless, the reaction mixture was quenched with saturated aqueous ammonium chloride (NH_4Cl) and was extracted with EtOAc. NMR spectroscopy confirmed that the reaction had failed to proceed and that starting material had been recovered.

A suggested reason for the result discussed previously was that an unprotected indole nitrogen has a proton that may be picked up by the ylide, thereby deactivating this reagent. It was also proposed that by introducing an electron-withdrawing protecting group on the indole nitrogen, the carbonyl carbon would be more electrophilic, thereby facilitating attack by the ylide on this functional group. To this end, the Boc-protected indole starting material, 1-(*tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (**42**), was synthesised. This was achieved by the addition of di-*tert*-butyl dicarbonate (Boc anhydride) to a solution of 3-benzoyl-5-chloro-1*H*-indole-2-carboxylate (**41**) and catalytic amounts of 4-dimethylaminopyridine (DMAP) in THF. The desired product was obtained after work-up and purification in 90% yield after a reaction of just 30 minutes at 30 °C.

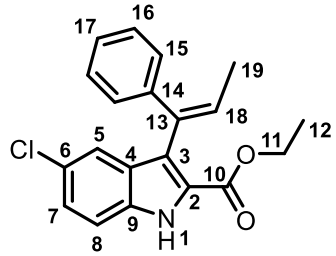


42

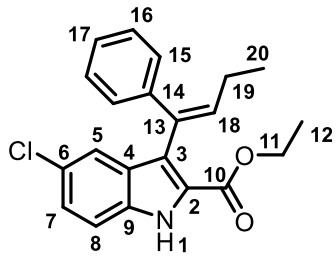
In terms of spectroscopic evaluation, the ^{1}H NMR spectroscopy signal at 1.63 ppm, integrating for 9 protons, as well as the ^{13}C NMR signals at 148.4 ppm for C_{18}, 86.6 ppm for C_{19} and 27.9 ppm for C_{20}, indicated the presence of the protecting group. In addition, the absence of a singlet in the ^{1}H NMR spectrum at 9.42 ppm confirmed that the protecting group had been introduced to the indole nitrogen atom. Finally, the NMR spectra also corresponded well to those reported in the literature.[1,9,17]



16



20



24

The Wittig reaction was then attempted a second time, this time on 1-(*tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (**42**). Satisfyingly, the reaction was carried out successfully, forming the desired alkenes with concomitant Boc deprotection, resulting in products obtained in 73% yield for the vinyl derivative (**16**), 51% yield for the propenyl derivative (**20**) and 41% yield for the butenyl derivative (**24**) respectively. It was noted that as the number of carbons in the Wittig reagent increased, the yield decreased. The unexpected deprotection was inconsequential and ultimately decreased the number of steps in the synthetic route.

The alkenylated products' structures were each confirmed by NMR spectroscopy. In all cases, the C=O signal of the ketone, in the ^{13}C NMR spectrum of starting material **42**, was absent in that of the products. In product **16**, the H_{18} protons were observed as two doublets, each integrating for 1 proton, at 5.99 ppm and 5.41 ppm. In product **20**, H_{18} was observed as two quartets belonging to the major and minor isomers of the product and integrating cumulatively for 1 proton at 6.44 ppm (major isomer) and 5.99 ppm (minor isomer) (*Figure 2.11*, circled in blue). Similarly, H_{18} was observed as two triplets at 6.35 ppm (major isomer) and 5.86 ppm (minor isomer) in **24**. Note that the major and minor isomers were not identified as *E* or *Z*, but that molecular modelling did not show any significant difference in binding energies between the two isomers. Furthermore, all carbon and proton atoms were accounted for in each product and mass spectrometry (MS) provided additional evidence of the structures synthesised. The NH peaks observed in the ^{1}H NMR spectra in each case, indicated deprotected products.

It was noted that a higher concentration of *n*-BuLi produced a higher yield in the Wittig reaction. For example, when 1.1 M *n*-BuLi (3 eq.) was used in the reaction to form the vinyl derivative **16**, a 47% yield was obtained compared to the 73% yield obtained when 1.88 M *n*-BuLi (3 eq.) was used. Concentrations of *n*-BuLi of 1.8 M or higher were thus found to be ideal to perform the alkenylation reactions. Note that, for the formation of compounds **20** and **24**, only the highest yields obtained are reported and that, in these cases, the concentrations of *n*-BuLi used ranged between 1.88 M and 1.89 M.

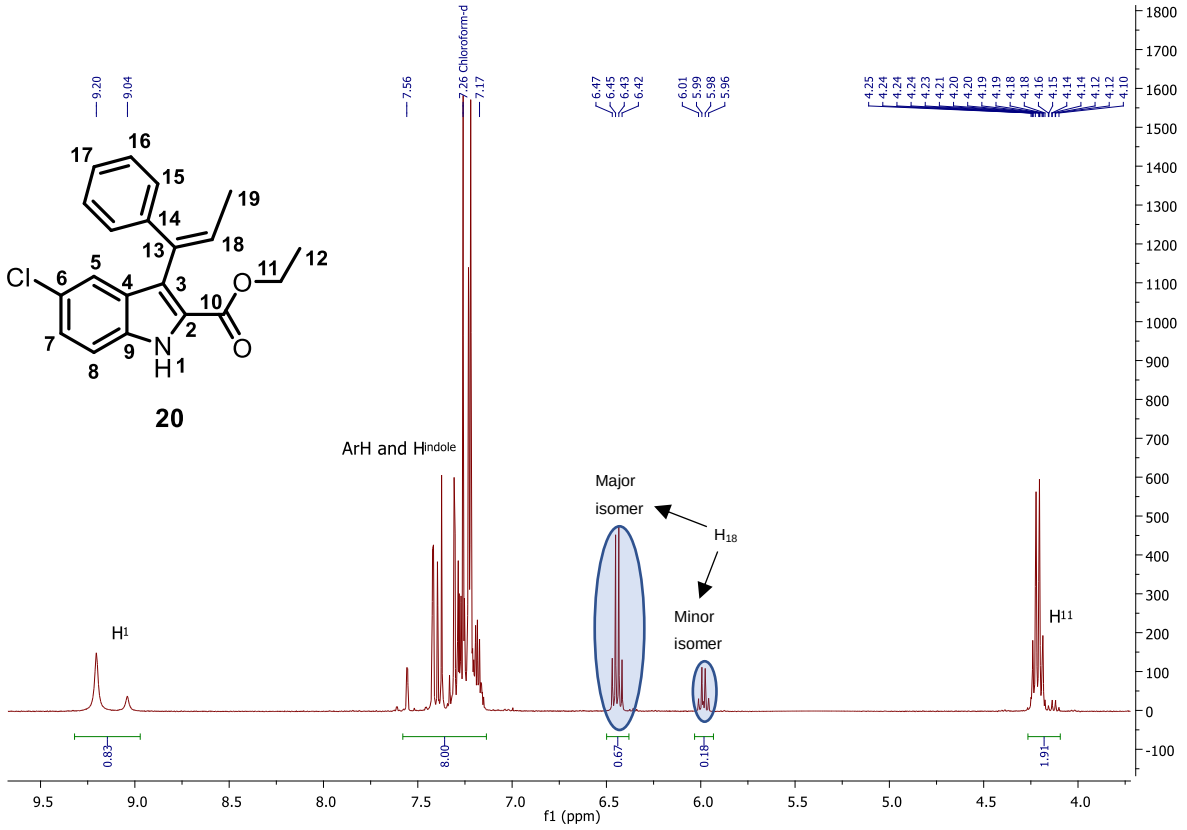


*Figure 2.11: The ^{1}H NMR spectrum of product **20** shown from 9.7 ppm to 3.7 ppm. From left to right the signals correspond to H_1 (major and minor isomers), ArH and H_{indole}, H_{18} (major and minor isomers) and H_{11} (major and minor isomers). The major and minor isomers of H_{18} are circled in blue.*

2.2.4 Hydrogenation of the alkenes

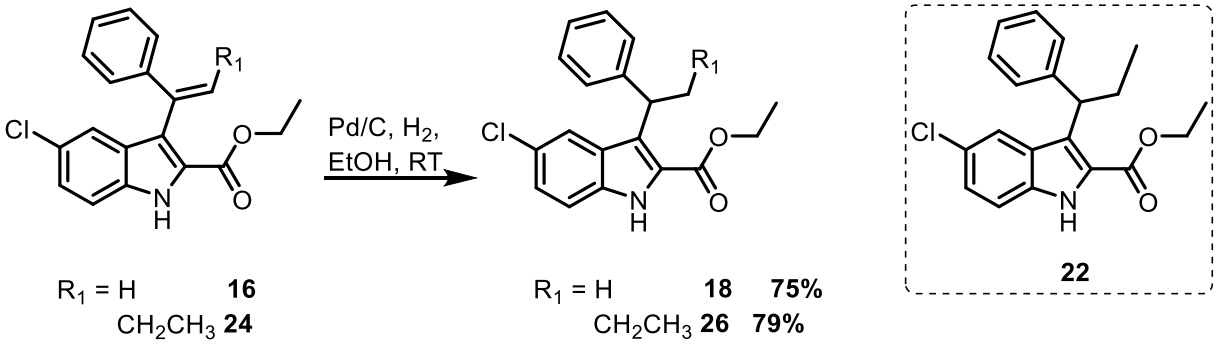
Paul Sabatier, a French Chemist in the late 1800s and the early 1900s, who has been described as the father of the chemical theory of catalysis, stated that during catalysis, an unstable intermediate, between the metal catalyst and one of the reactants, forms on the surface of the catalyst.[28,29] He observed that nickel was necessary for combining acetylene and hydrogen and concluded that the catalytic cycle started with nickel, or any metal catalyst, attracting the hydrogen onto its surface. This capricious hydrogen then moved from the metal and bonded with acetylene.[28] In 1912, Sabatier was awarded the Nobel Prize in chemistry for his work on the hydrogenation of organic compounds in the presence of finely divided metals.[28,29]

Catalysis has become important for the production of commercial chemicals, such as petrochemicals, synthetic fuels and fat hydrogenation. A particularly interesting application is the production of synthetic water, using carbon dioxide and hydrogen in the presence of a catalyst, on board the International Space Station.[28]

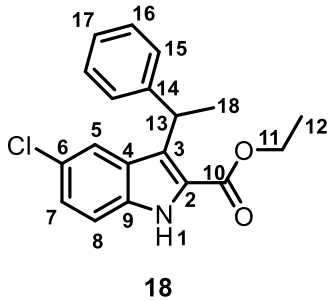
This simple and efficient means of hydrogenation was next implemented in the synthetic pathway to reduce the olefins formed in the previous step. The catalyst chosen was palladium absorbed onto carbon and a balloon filled with hydrogen gas was used to create the hydrogen atmosphere required.

2.2.4.1 Synthesis of ethyl 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylate (18) and ethyl 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylate (26)

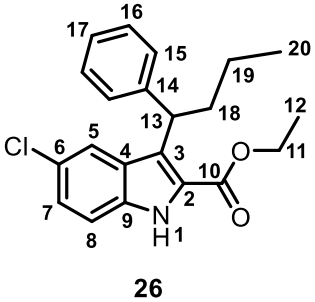
To the indole carboxylate dissolved in ethanol (EtOH), a spatula tip of the chosen metal catalyst, palladium on carbon (Pd/C), was added and the flask was placed under hydrogen (H_2) atmosphere. The reaction had to be monitored carefully by TLC, since dechlorinated products formed, if the reaction was left for too long, and a combination of the desired product and the starting material was obtained if the reaction ran for an insufficient amount of time. It was not possible to separate the starting material from the product by column chromatography, since the R_f values were nearly identical. The reaction thus had to be monitored until a small amount of the dechlorinated product was detected on the TLC plate. At this point, it was found that all the starting material had been consumed and the desired product could be separated from the dechlorinated product by column chromatography. Using this method, both ethyl 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylate (**18**) (75% yield) and ethyl 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylate (**26**) (79% yield) were successfully obtained from **16** and **24** respectively (*Scheme 2.6*). Ethyl 5-chloro-3-(1-phenylpropyl)-1*H*-indole-2-carboxylate (**22**) was synthesised by Ms Siobhan Brigg during her MSc project and was included in the biological evaluation for comparative purposes.[30]



*Scheme 2.6: The synthetic method used to synthesise compounds **18** and **26** from **16** and **24** respectively.*



18



26

In terms of spectroscopic evaluation, the C_{13} signals in the ^{13}C NMR spectra of both products shifted from the region between 140 ppm and 130 ppm to 5.36 ppm in **18** and 5.25 ppm in **26** respectively, indicating that the double bond had been reduced in each case. Furthermore, the C_{18} signals of compounds **18** and **26** were now located at 20.1 ppm in compound **18** and at 36.3 ppm in compound **26**. H_{18} presented as a doublet, integrating for 3 protons, in the ^1H NMR spectrum of **18**, while H_{18}

was a multiplet, integrating for 2 protons, in **26**. High resolution mass spectrometry (HRMS) was also used to confirm the presence of the $[M+H]^+$ ions in each case.

2.2.5 An amidation reaction

It was next proposed that the amide derivatives **17**, **19**, **21**, **25** and **27** (*Figure 2.12*) could be formed by firstly hydrolysing the esters into their corresponding carboxylic acids and then, by employing a well-known method, the acid chloride derivatives could be formed and subsequently converted into their corresponding amides by quenching with ammonium hydroxide solution.[18,30] It should be noted that Ms Siobhan Brigg synthesised 5-chloro-3-(1-phenylpropyl)-1*H*-indole-2-carboxamide (**23**) during her MSc project and that this compound was included in the biological evaluation for comparative purposes.[30]

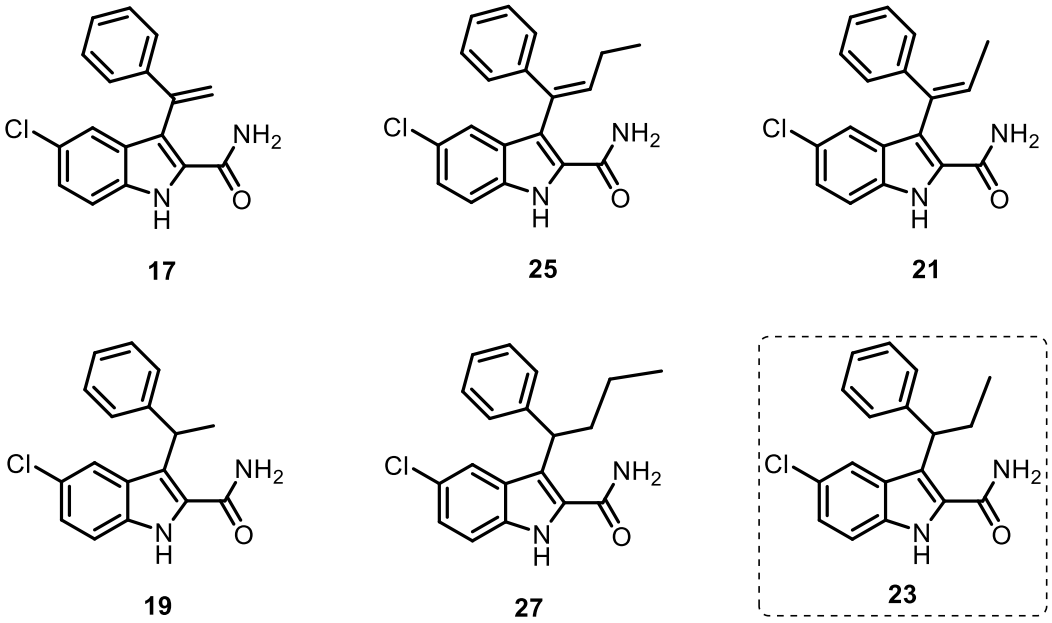
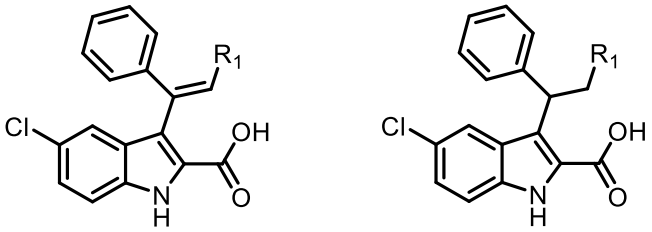


Figure 2.12: Anticipated amide derivatives.

2.2.5.1 Ester hydrolysis

For the ester hydrolysis, the indole carboxylate compounds (either **16**, **18**, **20**, **24** or **26**, Sections 2.2.3.1 and 2.2.4.1) were added to a round-bottom flask and were dissolved in a mixture of EtOH and water. Lithium hydroxide (LiOH) was then added and the reaction mixture was heated at reflux for 2 hours. After work-up, the corresponding crude carboxylic acid products (**43**, **44**, **45**, **46, 47**, *Figure 2.13*) were used as is in the next reactions. It should be noted that characterisation of these compounds was not done, but that successful amide formation (Section 2.2.5.2) indicated the successful conversion of the ester derivatives into their corresponding carboxylic acid derivatives.

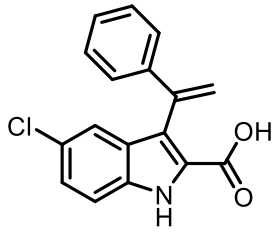


R_1 = H **43**
CH_3 **45**
CH_2CH_3 **46**

R_1 = H **44**
CH_2CH_3 **47**

*Figure 2.13: The resultant carboxylic acid derivatives, **43**, **44**, **45**, **46** and **47**, of indole carboxylate compounds **16**, **18**, **20**, **24** and **26** respectively.*

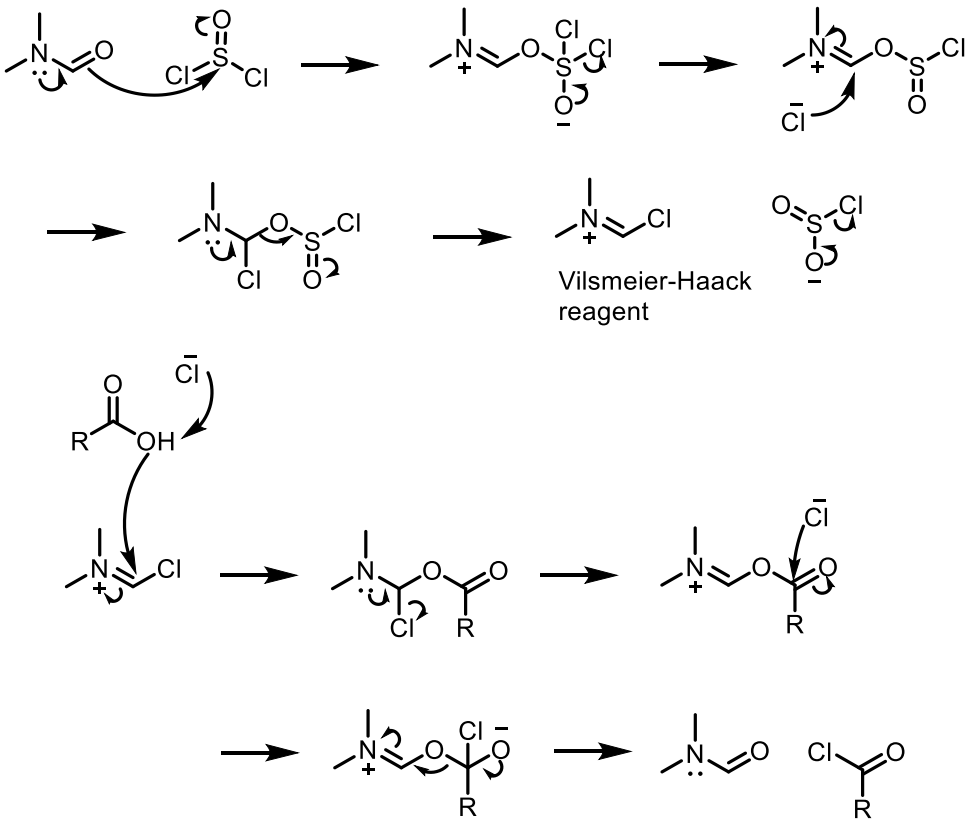
2.2.5.2 Amide formation[1,9,17,18,30]



43

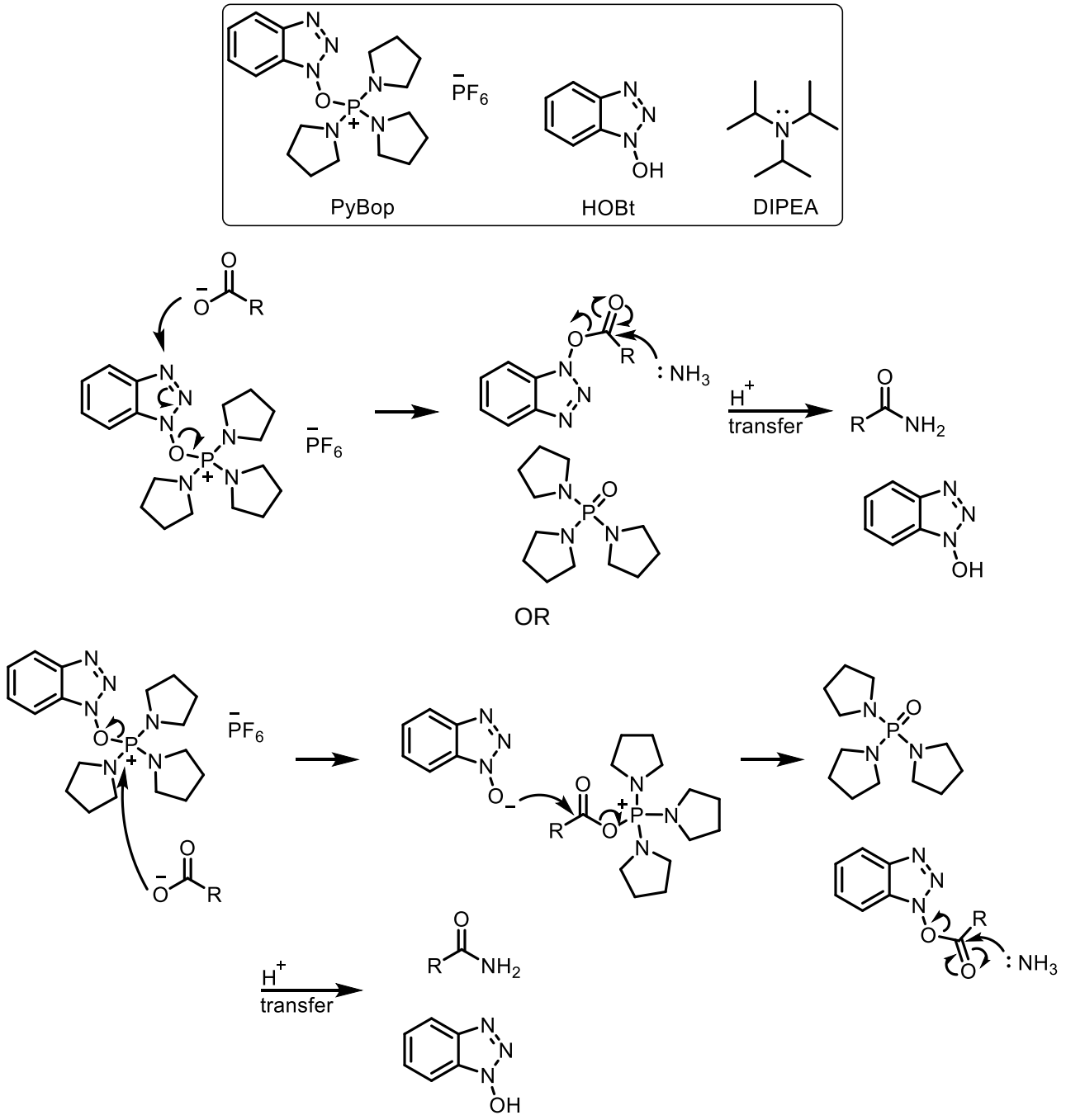
The next step involved the formation of the acid chlorides. To this end, carboxylic acid **43** and thionyl chloride ($SOCl_2$) were initially dissolved in chloroform (CH_3Cl). A catalytic amount of dimethylformamide (DMF) was added to facilitate the formation of the Vilsmeier-Haack reagent, which would ultimately react with the carboxylic acid, after which attack by a chloride ion would result in the acid chloride (*Scheme 2.7*). The reaction mixture was heated at reflux for 18 hours to ensure complete formation of the acid chloride. Next, amide formation was anticipated through a simple displacement of the chlorine atom with ammonia. To achieve this, the cooled reaction mixture was added dropwise to a beaker charged with 25% aqueous ammonium hydroxide solution. During addition, a low temperature was maintained by means of an ice-water bath. The solution was then left to stir on ice for 2 hours, after which extraction with EtOAc was done.

Although this is a well-known method for amide formation[17,18,30], it was not successful in this case. It is possible that the water in the ammonium hydroxide solution reversed the acid chloride formation. Upon TLC analysis, at least seven spots were detected and the material collected after column chromatography was not sufficient for NMR spectroscopic analysis.



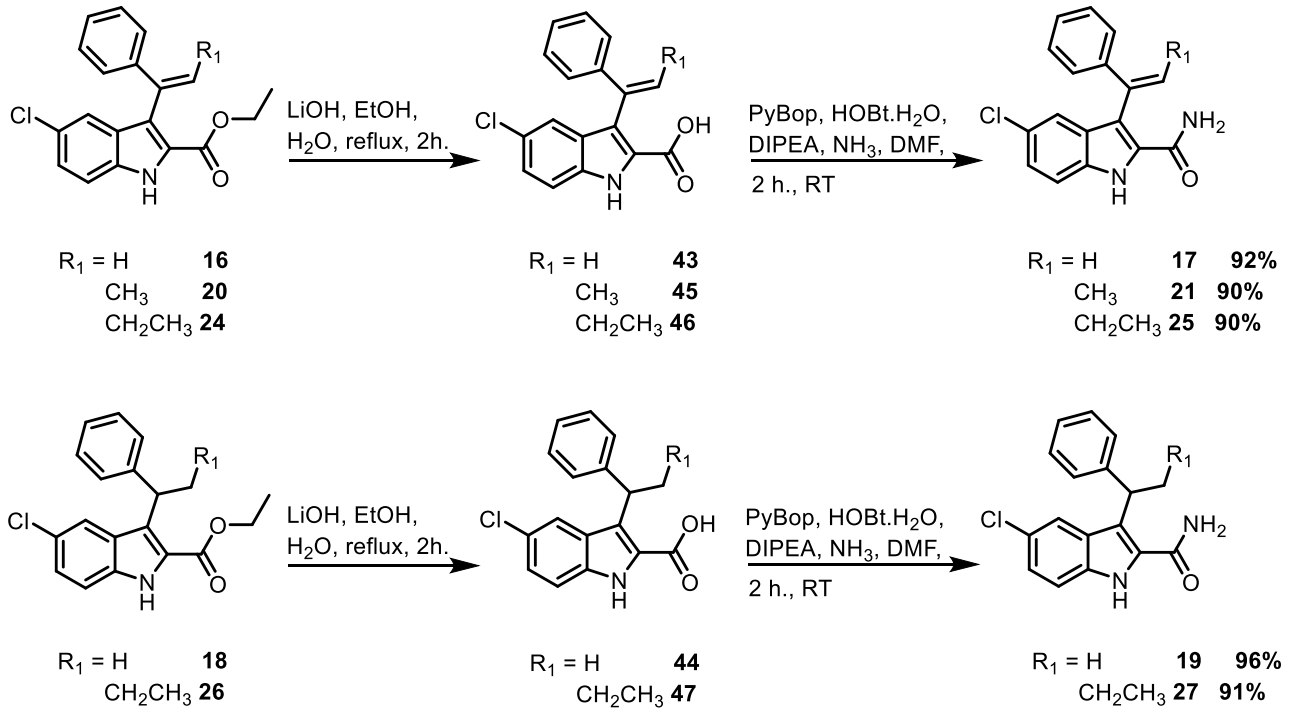
Scheme 2.7: DMF as catalyst in the formation of an acid chloride.

An alternative method for amidation involving the activation of the carboxylic acid moiety in the presence of the coupling reagent, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) (*Scheme 2.8*), was then attempted.[1] In this approach, during activation, the hydroxyl groups of the carboxylic acids (either **43**, **44**, **45**, **46** or **47**, *Figure 2.13*) were replaced with leaving groups (*Scheme 2.8*). Many other coupling agents are available, each with specific advantages. These are discussed in a comprehensive review by Valeur and Bradley.[31] Additives are used along with coupling reagents with the purpose of inhibiting side reactions and reducing racemization. The chosen additive in this case was 1-hydroxybenzotriazole hydrate (HOBt·H_2O) (*Scheme 2.8*).[1] DMF is necessary to solubilise HOBt·H_2O and is often used as the sole solvent for convenience. For the reaction to proceed, the carboxyl groups of **43**, **44**, **45**, **46** or **47** (*Figure 2.13*) must be in anionic form, thus the addition of the tertiary amine, *N*,*N*-diisopropylethylamine (DIPEA) (*Scheme 2.8*).[1]

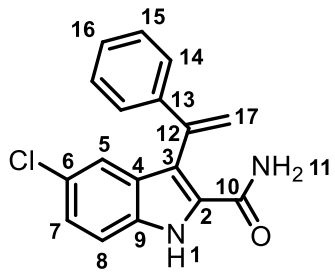


Scheme 2.8: Two possible reaction mechanisms for PyBop-mediated amide formation.

The reactions proceeded when ammonia gas was bubbled through solutions of the carboxylic acid derivatives (either **43**, **44**, **45**, **46** or **47**, *Scheme 2.9*), PyBop, HOBt·H_2O and DIPEA in DMF. After stirring at room temperature for 2 hours under ammonia atmosphere, general work-up procedures were followed and purification by column chromatography resulted in the pure desired products (5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (**17**), 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxamide (**19**), 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxamide (**21**), 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxamide (**25**) and 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxamide (**27**) respectively) (*Scheme 2.9*).



*Scheme 2.9: A summary of the ester hydrolysis and amidation reactions used to synthesise final compounds **17**, **19**, **21**, **25** and **27**.*



17

The final products (after two steps) were each confirmed with ^{1}H and ^{13}C NMR spectroscopy, as well as with HRMS. An obvious indication of product formation was the absence of the ethyl chains of the ester functionalities in the indole carboxylate starting materials (**16**, **18**, **20**, **24** and **26**, *Scheme 2.9*). In all cases, the two protons of the amide moiety were accounted for. The ^{1}H NMR spectrum of **17** is shown as an example in *Figure 2.14*. Here the two protons of the amide moiety are clearly visible as two separate broad singlets at 6.56 ppm and 5.89 ppm (circled in green).

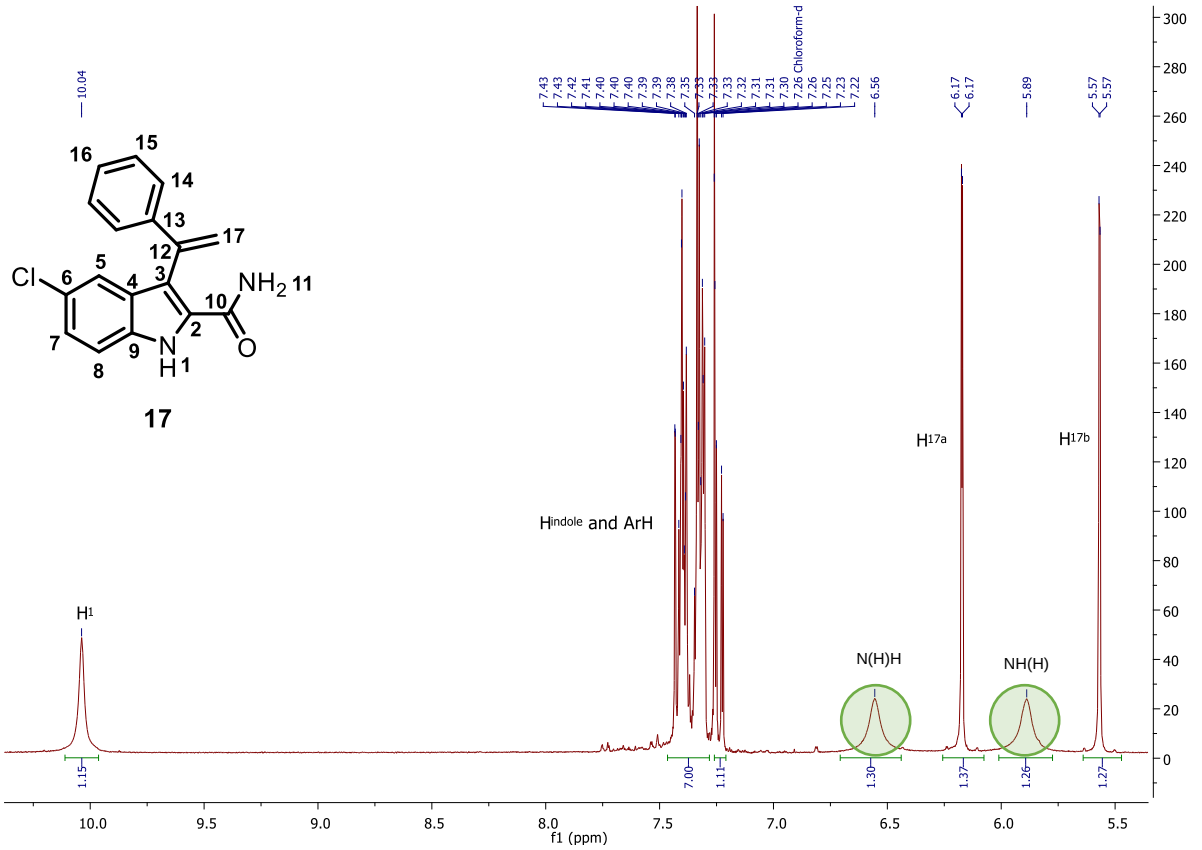


*Figure 2.14: The ^{1}H NMR spectrum of product **17** shown from 10.4 ppm to 5.4 ppm. From left to right the signals correspond to H_1, ArH and H_{indole}, N(H)H, H_{17a}, NH(H) and H_{17b} as labelled. The amide proton signals are circled in green.*

2.2.6 Biological results and concluding remarks

To conclude this section, biological evaluation was carried out on all final compounds, listed in *Table 2.4*, by our collaborator, Dr. Adriaan Basson, at the HIV Pathogenesis Research Unit at the Medical School, University of the Witwatersrand. An *in vitro* single-cycle, non-replicative phenotypic assay, using a HIV-1 retroviral vector system for the production of virus-like particles, was utilised.[1] To this end, the compounds were incubated with 293T cells together with the virus-like particles for 48 hours, after which inhibition was quantified by luminescence measurement. Briefly, 100 μL of BrightGlo luciferase substrate (Promega) was added to the culture wells. After incubation for 3 minutes at room temperature, the contents of the wells were homogenized with a multichannel pipette and transferred to the wells of a white 96-well plate. Luminescence in the wells of the white 96-well plate was read on a GloMax Explorer system, with a 0.5 second read per well. The toxicity and activity assay results are presented in *Table 2.4* below. It should be noted that nevirapine had an IC_{50} value within the expected range and was used as a control.

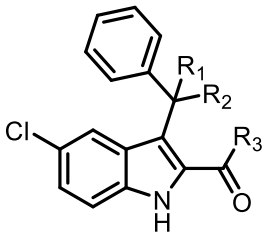
Firstly, six compounds (highlighted in green, *Table 2.4*), two of which were synthesised by Siobhan Brigg during her MSc project[30], were found to be more potent than nevirapine against wild-type HIV-1. Four of these, compounds **20** – **23**, had a chain of two carbon atoms in the benzylic position.

From these results, the optimum chain length for occupying the Val179 pocket was inferred to be two carbon atoms.

Next, the vinyl (**16** and **17**, *Table 2.4*), propenyl (**20** and **21**) and butenyl (**24** and **25**) derivatives were compared to their ethyl (**18** and **19**), propyl (**22** and **23**) and butyl (**26** and **27**) counterparts. While compounds **18** and **19** proved to be more potent than their $=CH_2$ counterparts, compounds **20**, **21**, **22** and **23** all showed great potency. No trend was observed for compounds **24** – **27**.

Lastly, the effect of an amide group compared to an ester group, in the indole 2-position, on efficacy was considered. For compounds with an unsaturated chain in the benzylic position (**16**, **17**, **20**, **21**, **24**, **25**, *Table 2.4*), the amide derivatives had a higher efficacy than the ester derivatives, while the opposite was seen for compounds with a saturated chain in the benzylic position (**18**, **19**, **22**, **23**, **26**, **27**, *Table 2.4*). Comparing the amide and ester derivatives of compounds with a two-carbon chain in the benzylic position (**20** – **23**, *Table 2.4*), it was noted that the IC_{50} values were very similar. This points towards the importance of the chain length in the benzylic position for efficacy rather than the functional group in the indole 2-position.

Table 2.4: A summary of the toxicity and activity data obtained from biological tests against wild-type HIV-1 for all final compounds synthesised in Section 2.2.



Compounds			Toxicity (CC_{50}, μM)				Activity (IC_{50}, μM)			
#	R_1, R_2	R_3	1	2	Ave	SD	1	2	Ave	SD
16	$=CH_2$	–OEt	64.81	70.05	67.43	3.71	0.397	0.444	0.421	0.034
17	$=CH_2$	$–NH_2$	45.42	37.03	41.22	5.93	0.325	0.332	0.329	0.005
18	$–CH_3$, H	–OEt	57.27	56.45	56.86	0.58	0.010	0.016	0.013	0.004
19	$–CH_3$, H	$–NH_2$	43.51	39.80	41.65	2.63	0.144	0.124	0.134	0.013
20	$=CH–CH_3$	–OEt	53.23	50.25	51.24	2.10	0.064	0.050	0.057	0.010
21	$=CH–CH_3$	$–NH_2$	30.51	23.55	27.03	4.92	0.009	0.011	0.010	0.001
22*	$–CH_2–CH_3$, H	–OEt	52.16	53.09	52.62	0.66	0.016	0.015	0.016	0.000
23*	$–CH_2–CH_3$, H	$–NH_2$	24.09	23.48	23.78	0.43	0.020	0.017	0.019	0.002
24	$=CH–CH_2–CH_3$	–OEt	51.23	54.18	52.71	2.09	0.366	0.371	0.369	0.004
25	$=CH–CH_2–CH_3$	$–NH_2$	28.47	24.68	26.58	2.67	0.202	0.213	0.208	0.008
26	$–CH_2–CH_2–CH_3$, H	–OEt	55.74	55.60	55.67	0.10	0.086	0.078	0.082	0.006
27	$–CH_2–CH_2–CH_3$, H	$–NH_2$	22.46	21.46	21.95	0.70	0.327	0.309	0.318	0.013
NVP							0.105	0.076	0.091	0.020

Ave = average

SD = standard deviation

NVP = Nevirapine control

* = compounds not synthesised during this MSc project.

Green highlight = compounds that are more potent than nevirapine against wild-type HIV-1.

These promising efficacy results provided valuable information for future studies regarding the optimum chain length for occupying the Val179 pocket (two carbon chain) and the tolerance for various groups in the indole 2-position.

2.3 More adventurous amides

2.3.1 A known synthetic route

A synthetic route to obtain the indole sulphide **54** has been well established and was reported by S. Brigg *et al. in 2016* (*Scheme 2.10*).[10] In terms of a brief summary, the commercially available ethyl

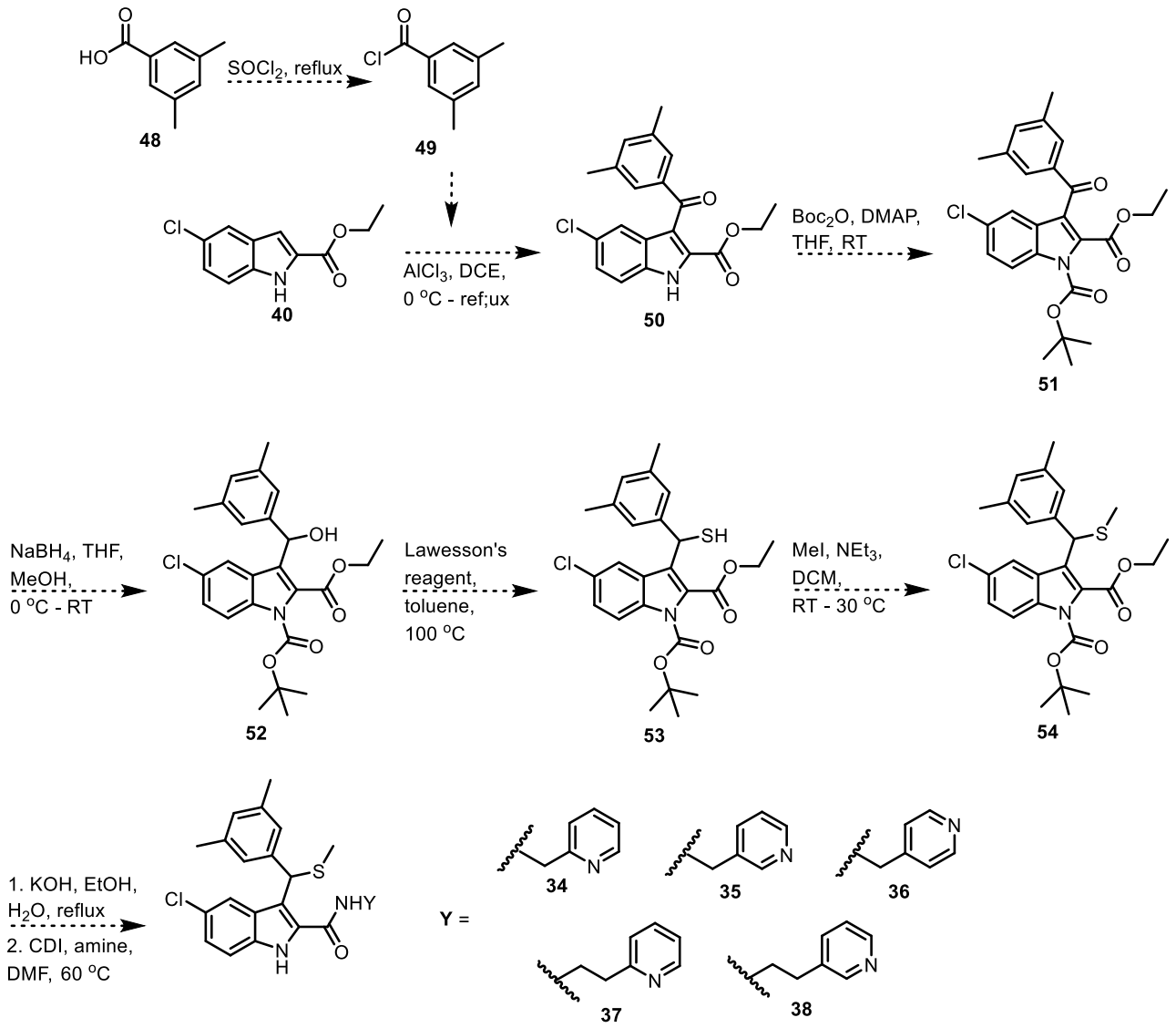
5-chloro-2-indole-carboxylate (**40**) was acylated by means of a Friedel-Crafts acylation reaction to obtain **50**. Subsequent Boc-protection (**51**), followed by ketone reduction with sodium borohydride readily afforded **52**. Alcohol **52** was then treated with Lawesson's reagent and the corresponding thiol **53** was methylated to successfully obtain **54**. It was envisaged that the ester moiety of **54** could be hydrolysed with a suitable base and that a coupling reaction could afford the desired amides. The next section describes the synthesis of these compounds.

2.3.2 Another Friedel-Crafts acylation reaction

For the introduction of two methyl groups at positions 3 and 5 of the 3-[(phenyl)(methylthio)] moiety, important for potency against resistant strains of HIV, 3,5-dimethylbenzoyl chloride (**49**) was required as a reagent.[32] Since 3,5-dimethylbenzoyl chloride (**49**) is a costly reagent needed in a large quantity, it was synthesised from available 3,5-dimethylbenzoic acid (**48**), rather than being purchased.

2.3.2.1 Synthesis of 3,5-dimethylbenzoyl chloride (49)[10,30]

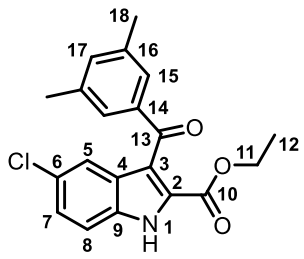
3,5-Dimethylbenzoic acid (**48**, *Scheme 2.10*) was thus heated at reflux in the presence of thionyl chloride to form the corresponding acid chloride (**49**). A long reaction time of 24 hours was used to ensure for optimum conversion. The excess thionyl chloride was then removed by heating under vacuum using a trolley vacuum pump equipped with a base trap. The brown oil was used directly in the Friedel-Crafts reaction without purification.



Scheme 2.10: A synthetic strategy towards the target indole amides.

2.3.2.2 Synthesis of ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-2-carboxylate (50)[9,10,30]

3,5-Dimethylbenzoyl chloride (**49**) was then transferred to a round-bottom flask containing dry DCE, while taking care to avoid moisture contamination, since this would reverse the acid chloride formation. The solution was cooled using an ice-water bath before $AlCl_3$ was added. To allow for the formation of the acylium ion, the reaction mixture was stirred for 30 minutes prior to the addition of ethyl-5-chloro-1*H*-indole-2-carboxylate (**40**). The reaction mixture was then heated at reflux for 3 hours, followed by a work-up and product purification.



50

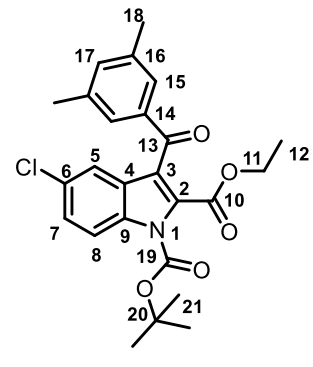
The successful formation of the title product **50** was confirmed using NMR spectroscopy and the spectra compared well to those reported in the literature.[9,10,30] Perhaps the easiest way to recognise the product was to identify the expected large singlet corresponding to H_{18} and the triplet corresponding to H_{12} in the ^{1}H NMR spectrum of the product. An integration ratio could then be calculated. This should be 2:1 since the singlet should integrate for 6 protons (H_{18}, 2 x ArMe) and the triplet for three (H_{12}, 3H). Indeed, this was the case. A second easily recognisable peak corresponded to C_{13} and was found at 193.0 ppm in the ^{13}C NMR spectrum.

2.3.3 Installing a protecting group

Previous studies on the indole scaffold showed that a protecting group on the indole nitrogen was key to obtain optimum yields.[9,10,17,30] Furthermore, methylation (Section 2.3.6) without a protecting group would result in a methyl substituent on the indole nitrogen. The *tert*-butyloxycarbonyl (Boc) protecting group seemed suitable for use in this synthetic route, since this protecting group is commonly used in syntheses where Lawesson's reagent is employed, particularly because the Boc group is heat stable and high temperatures are typically required for the Lawesson's reaction.[33] The Boc group is also easily removable using a base and was expected to be removed upon ester hydrolysis.

2.3.3.1 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-1,2-dicarboxylate (51)[9,10,30]

For the synthesis of compound **51**, a typical synthetic method for installing the Boc group was used, as was described in Section 2.2.3.1. The reaction was carried out in THF at room temperature using DMAP (catalytic) and Boc_2O. This reaction proved to be highly effective, with the starting material consumed after 30 minutes and a 91% yield being obtained after work-up and purification by column chromatography.



51

In terms of characterisation by NMR spectroscopy, a large singlet at 1.63 ppm in the ^{1}H NMR spectrum of the product, which integrated for 9 protons, represented the newly introduced tertiary butyl group. Furthermore, C_{19}, C_{20} and C_{21} were identified in the ^{13}C NMR spectrum at 148.5 ppm, 86.5 ppm and 27.9 ppm respectively. In addition, the absence of a singlet integrating for 1 proton at 9.66 ppm in the ^{1}H NMR spectrum, was proof of protection at the indole nitrogen. The NMR spectra corresponded well to those reported in the literature.[9,10,30]

2.3.4 A selective ketone reduction

In his Nobel Lecture, H. C. Brown told the interesting story of the discovery of sodium borohydride by himself and his boss, Schlesinger, in 1942 at the University of Chicago. The journey started when Brown's girlfriend, Sarah Baylen, presented him with a graduation gift when he received his B.S. degree from the University of Chicago in 1936. The gift, Alfred Stock's book, *The Hydrides of Boron and Silicon*, was the most economical chemistry book ($2.60) Sarah could find in the time of the Great Depression. In the words of Brown "Such are the developments that can shape a career!".[34]

After completing his Ph.D. thesis in 1938, Brown became research assistant to Professor Schlesinger who worked on developing metal borohydrides for wartime applications. Specifically, volatile uranium compounds with low molecular weights were of interest.[34] Uranium borohydride [$U(BH_4)_4$] suited the requirements well; however, the synthesis of this compound required the use of lithium hydride, a reagent that was in very short supply at that time. As a result, sodium hydride was identified as a possible substitute. Due to the poor solubility of sodium hydride in the solvents that were available, sodium trimethoxyborohydride was synthesised to solve the problem.[34,35]

$$NaH + B(OCH_3)_3 \rightarrow NaBH(OCH_3)_3$$

It was then found that diborane displaced methyl borate from sodium trimethoxyborohydride. As a result, sodium borohydride was prepared for the first time.[35]

$$2NaBH(OCH_3)_3 + (BH_3)_2 \rightarrow 2NaBH_4 + 2B(OCH_3)_3$$

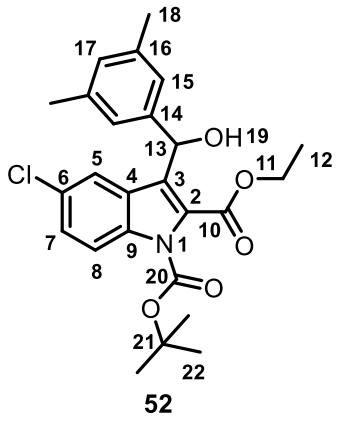
When $NaBH_4$ was dissolved in acetone, rapid reduction of acetone to the corresponding alcohol was observed. In this manner, it was discovered that sodium borohydride was a valuable reagent for the reduction of organic molecules.[34]

$$NaBH_4 + 4R_2C{=}O \rightarrow NaB(OCHR_2)_4 \rightarrow NaB(OH)_4 + R_2CHOH$$

Today, sodium borohydride is commonly used in laboratories and in industry as a mild reducing agent, of aldehydes, ketones and acid chlorides, with stable performance and selective reduction. This reagent was therefore chosen for our required selective reduction of a ketone in the presence of esters.

2.3.4.1 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (hydroxy)methyl]-1*H*-indole-1,2-dicarboxylate (52)[9,10,30]

The reduction reaction proceeded without difficulty when the protected indole, **51,** was added to round-bottom flask containing THF and MeOH and the solution was cooled to 0 °C, using an ice-water bath, before sodium borohydride ($NaBH_4$) was added. Next, the reaction mixture was warmed to room temperature. After 3 hours, the starting material was consumed, as per TLC analysis, and a work-up was done. Purification by column chromatography then afforded the pure title compound in 90% yield.



52

Upon analysis of the ^{1}H NMR spectrum of the product (*Figure 2.15*), H_{13} and H_{19} (circled in orange) were observed as two doublets, each integrating for 1 proton at 6.16 ppm and at 5.99 ppm. Reduction of the ketone to an alcohol was therefore successful. In the ^{13}C NMR spectrum, C_{10} and C_{20} were observed far downfield at 161.6 ppm and 148.4 ppm respectively, indicative of carbonyl carbons. The ketone had therefore been selectively reduced. Additionally, both the ^{13}C NMR spectrum and the ^{1}H NMR spectrum compared well to those reported in the literature.[9,10,30]

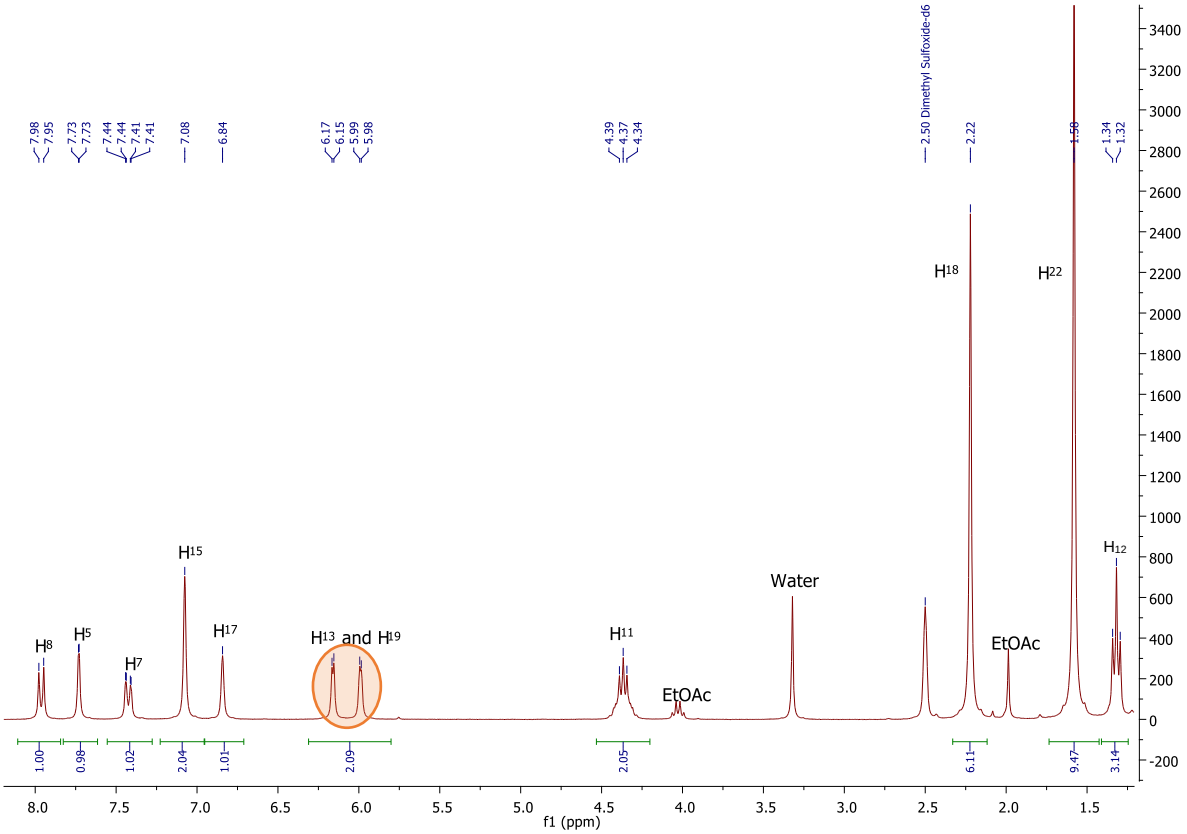


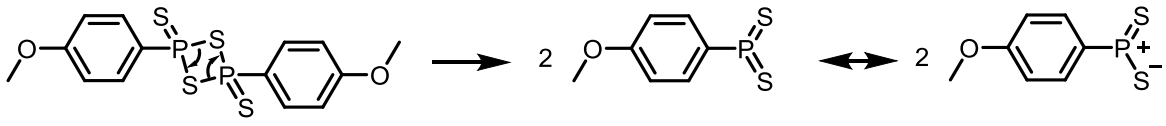
*Figure 2.15: The ^{1}H NMR spectrum of product **52** shown from 8.2 ppm to 1.2 ppm. From left to right the signals correspond to H_8, H_5, H_7, H_{15}, H_{17}, $H_{13/19}$, $H_{19/13}$, H_{11}, H_{18}, H_{22} and H_{12} as labelled. The proton signals corresponding to H_{13} and H_{19} are circled in orange.*

2.3.5 The use of Lawesson's reagent

The reagent 2,4-bis(4-methoxyphenyl)-1,3,2,4-dithiadiphosphetane 2,4-disulfide, although not initially invented by Lawesson, is commonly known as Lawesson's reagent after the scientist who studied the reactive properties of this reagent extensively, thereby making it a popular chemical reagent for thionation.[33,36] It is a versatile reagent that converts oxygen functionalities into their thio-analogs. Some selectivity exists in that a hydroxyl group will react before an amido group; similarly, an amido group will react before an ester group. In general terms, the oxygen functionality with the

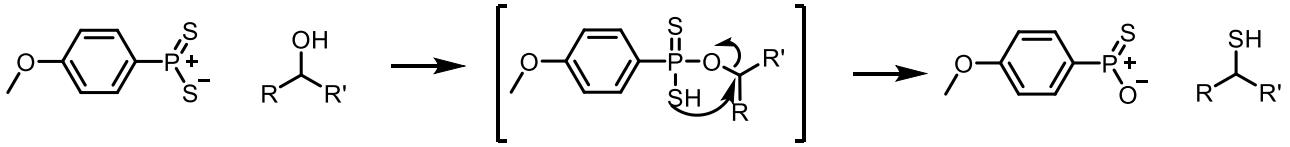
greater electron density is favoured in the reaction. Steric hindrance has also been seen to play a role in selectivity.[36]

It is believed that Lawesson's reagent breaks into two equivalent active species, (4-methoxyphenyl)phosphine disulfide (*Scheme 2.11*).[33]



Scheme 2.11: The dissociation mechanism of Lawesson's reagent.

The proposed mechanism of action for the formation of thiols starts with the reaction of an active species with the hydroxy group to form the intermediate shown in brackets in *Scheme 2.12*. This intermediate then decomposes to yield the thionated product.[37]

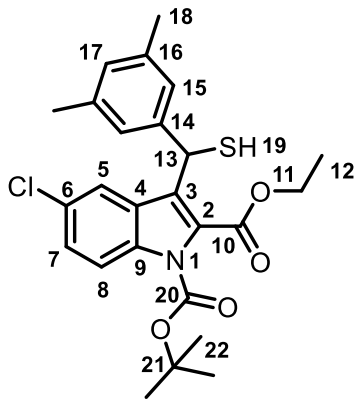


Scheme 2.12: The proposed thionation mechanism of Lawesson's reagent and alcohols.

The direct conversion of alcohols into thiols was achieved for the first time in 1993, through the use of Lawesson's reagent, by Takehiko Nishio at the University of Tsukuba, Japan.[37] A method adapted from his research was optimised for this specific indole system.[10,30]

2.3.5.1 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (mercapto)methyl]-1*H*-indole-1,2-dicarboxylate (53)[10,30]

First, 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(hydroxy)methyl]-1*H*-indole-1,2-dicarboxylate (**52**) was dissolved in toluene, after which Lawesson's reagent was added. The reaction mixture was then heated to 100 °C for 30 minutes. Although Nishio *et al*. reported that hydroxyl groups are the most reactive and ester groups the least among hydroxy, amide, ketone and ester functionalities, a lower than usual reaction temperature and a shorter time was used to reduce the probability of the two ester groups reacting.[33] After cooling the reaction mixture to room temperature, the solvent was removed under reduced pressure and the crude material was purified by column chromatography. A modest yield of 38% was achieved which was, however, sufficient to continue to the next step.



53

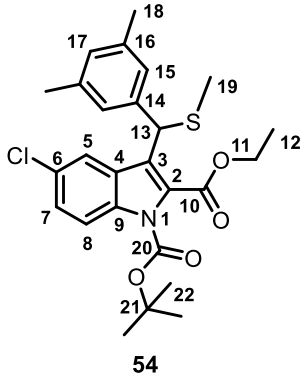
In terms of spectroscopic characterisation of **53**, the ^{1}H NMR spectrum of the product showed a shift of the doublets found at 6.19 ppm and 5.99 ppm for H_{13} and H_{19} in the starting material (**52**) spectrum, to 5.81 ppm and 2.37 ppm respectively. The lower electronegativity value of sulfur (2.5 compared to 3.5 of oxygen) increased the electron density around the proton. The proton was therefore more shielded and the chemical shift was expected to decrease, as was observed. Furthermore, the rest of the NMR spectroscopic data corresponded well to that in the literature.[10,30]

2.3.6 Methylation with iodomethane

Methyl iodide is a common alkylating agent for the methylation of carbon, oxygen, sulfur, nitrogen and phosphorus.[38] The sterically "open" methyl iodide is readily attacked by a nucleophile at the carbon atom, while the iodide, being a good leaving group, leaves simultaneously. An S_N2 substitution reaction therefore takes place.

2.3.6.1 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (methylthio)methyl]-1*H*-indole-1,2-dicarboxylate (54)[10,30]

A final alteration to the functional group in the benzylic position was the formation of a thioether. This was accomplished by first stirring 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(mercapto)methyl]-1*H*-indole-1,2-dicarboxylate (**53**) together with triethylamine (NEt_3) for 15 minutes in dichloromethane (DCM) at room temperature. The methylating agent, methyl iodide (MeI), was then added and the reaction mixture was stirred at 30 °C overnight. Since the starting material and the product had the same R_f value, a long reaction time (18 hours) was used for optimum conversion. After purification by column chromatography, minor starting material signals were still present in the NMR spectra. For this reason, a yield is not reported for this step and was calculated over three steps for the final products. In retrospect, adding a larger molar quantity of MeI than NEt_3 may have improved the yield, since there is the possibility of methylating NEt_3.



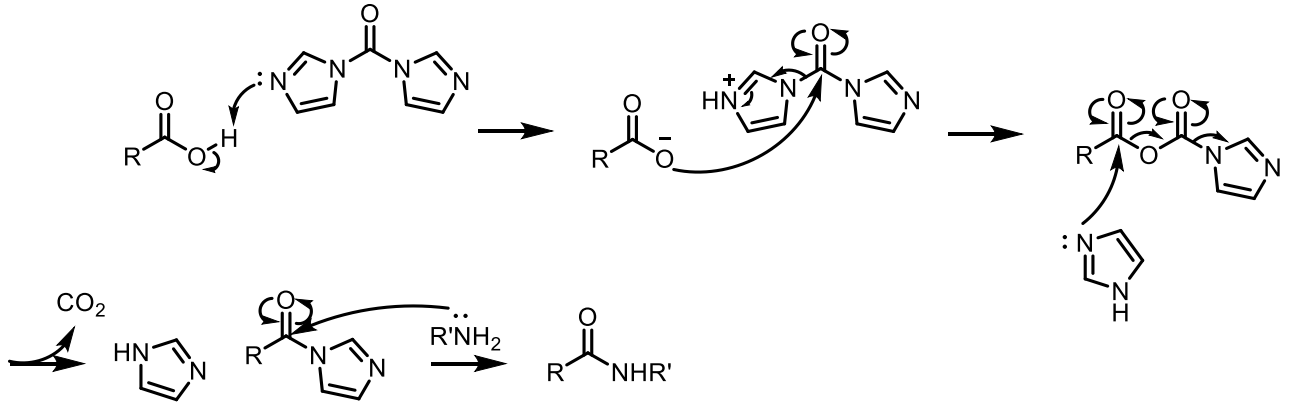
54

Nevertheless, the formation of the product **54** was confirmed from the NMR spectra. H_{19} was observed as a singlet, integrating for 3 protons, at 2.02 ppm in the ^{1}H NMR spectrum. In addition, C_{19} was observed at 16.0 ppm in the ^{13}C NMR spectrum. H_{13} was also no longer observed as a doublet, but rather as a singlet at 5.43 ppm in the ^{1}H NMR spectrum. These results compare well to data available in the literature.[10,30]

2.3.7 Amidation

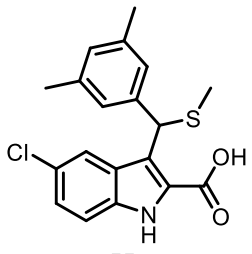
Carbonyl diimidazole (CDI) is a coupling reagent that allows for one-pot amide formation from a carboxylic acid.[39] This coupling reagent was inspired by Wieland and Schneider, who showed that peptide derivatives could be synthesised through acylation of the imidazole ring of methyl *N*-benzoyl-L-histidinate, followed by reaction with an appropriate amine.[40,41] Paul and Anderson then developed a more direct method for synthesising acyl-imidazoles using *N*,*N'*-carbonyldiimidazole.[41] They envisaged the utilisation of a more convenient reagent with the only by-products being imidazole and carbon dioxide. Furthermore, carbon dioxide evolution would drive the reaction forward.[41]

The CDI-mediated coupling reaction starts with the formation of the acylimidazole, generating the by-products, carbon dioxide and imidazole (*Scheme 2.13*). The desired amine which reacts at the carbonyl carbon, is then added. Imidazole, being a good leaving group, leaves.[39]



Scheme 2.13: A one-pot amide preparation using CDI.

2.3.7.1 Ester hydrolysis – Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (55)



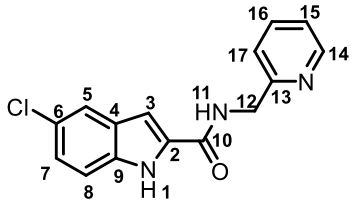
55

For the synthesis of a small library of amides, the carboxylic acid, 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**), was required. To this end, 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-1,2-dicarboxylate (**54**) was added to a round-bottom flask containing a mixture of EtOH and water. Potassium hydroxide (KOH) was then added and the flask was fitted with a reflux condenser. The reaction mixture was heated at reflux for 18 hours to ensure the completion of the reaction, before it was cooled to room temperature, acidified to a pH of 3 and extracted with EtOAc. No further purification was done to the solid that was isolated.

2.3.7.2 Test reaction – Synthesis of 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (56)

In order to test the planned synthetic method, the carboxylic acid, 5-chloro-1*H*-indole-2-carboxylic acid (kindly supplied by Ms Siobhan Brigg who synthesised this compound during her MSc project[30]),

was heated to 60 °C in DMF, after which 1,1'-carbonyldiimidazole (CDI) was added and the reaction mixture stirred for 3 hours to allow for the formation of the acylimidazole (*Scheme 2.11*). 2-Picolylamine was then added and the reaction mixture was stirred for a further 18 hours. Upon completion, the solvent was removed by heating under vacuum before the crude product was re-dissolved in EtOAc and loaded onto silica for purification by column chromatography. In this manner, the pure title compound was obtained in 75% yield.



56

The easiest way to recognise the product from the ^{1}H NMR spectrum was to identify H_{11} and H_{12}. These were observed as a triplet, integrating for 1 proton, at 9.20 ppm for H_{11} and a doublet, integrating for 2 protons, at 4.60 ppm for H_{12}. In addition, the newly incorporated pyridinyl aromatic protons were clearly observed in the ^{1}H NMR spectrum as a multiplet at 8.57 – 8.48 ppm (1H), a doublet of doublets of doublets at 7.77 ppm (1H), a doublet at 7.44 ppm (1H) and a multiplet at 7.31 – 7.24 ppm (1H). Finally, the correct number of carbon atoms and protons were present in the spectra and the $[M+H]^{+}$ ion was confirmed by HRMS (calculated 286.0747, found 286.0740).

2.3.7.3 The synthesis of the final amide products

In terms of the final targets in this section, 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**34**) was prepared in exactly the same manner as 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**) and a 29% yield was obtained over the three synthetic steps. The other amides (**35**, **36**, **37** and **38**) were prepared in a similar manner, but for practical purposes, THF was used as solvent instead of DMF, which has a very high boiling point and is difficult to remove. Yields ranging from 20 – 45% over three steps were obtained for these compounds (*Figure 2.16*). All product structures were confirmed with NMR spectroscopy and HRMS. The low yields obtained were possibly due to incomplete methylation (see Section 2.3.6), side reactions or low purity of the amines used. The amines were not purified after purchase due to the small quantities we had available.

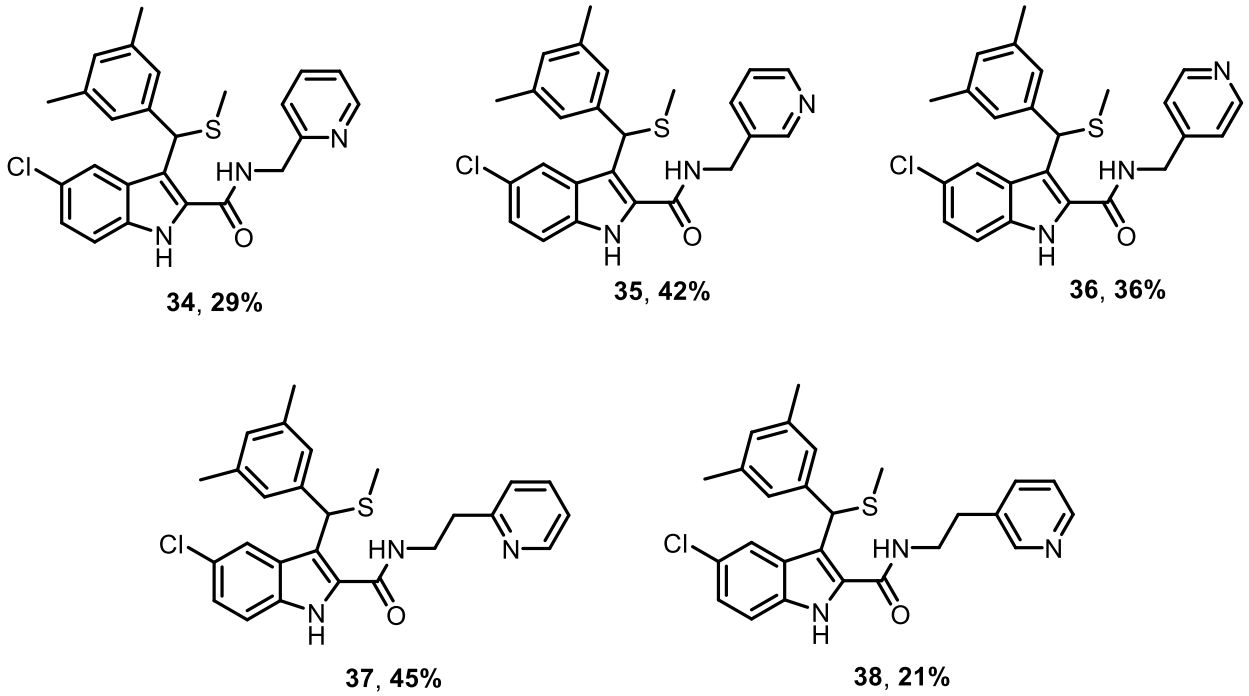
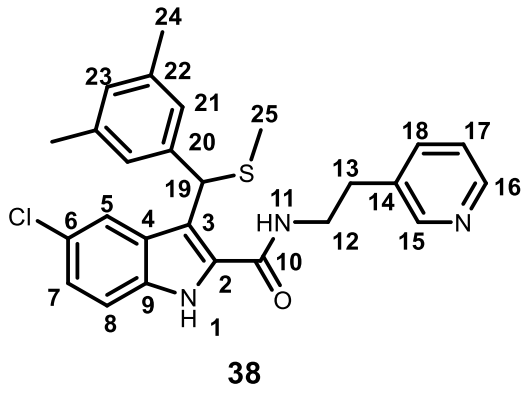


Figure 2.16: The final amide products and the yields obtained over three synthetic steps.



As a representative example, the ^{1}H NMR spectrum of compound **38** is presented in *Figure 2.17*. As with product **56**, protons H_{11}, H_{12} and H_{13} (circled in pink) proved to be key when determining the success of this reaction using ^{1}H NMR spectroscopy. In this case, H_{11} was observed as a triplet at 7.95 ppm. Perhaps more obvious are the signals of H_{12} and H_{13}, as these were detected as two multiplets at 3.77 – 3.52 ppm and 2.93 – 2.59 ppm. Finally, the pyridinyl protons were also evident in the spectrum.

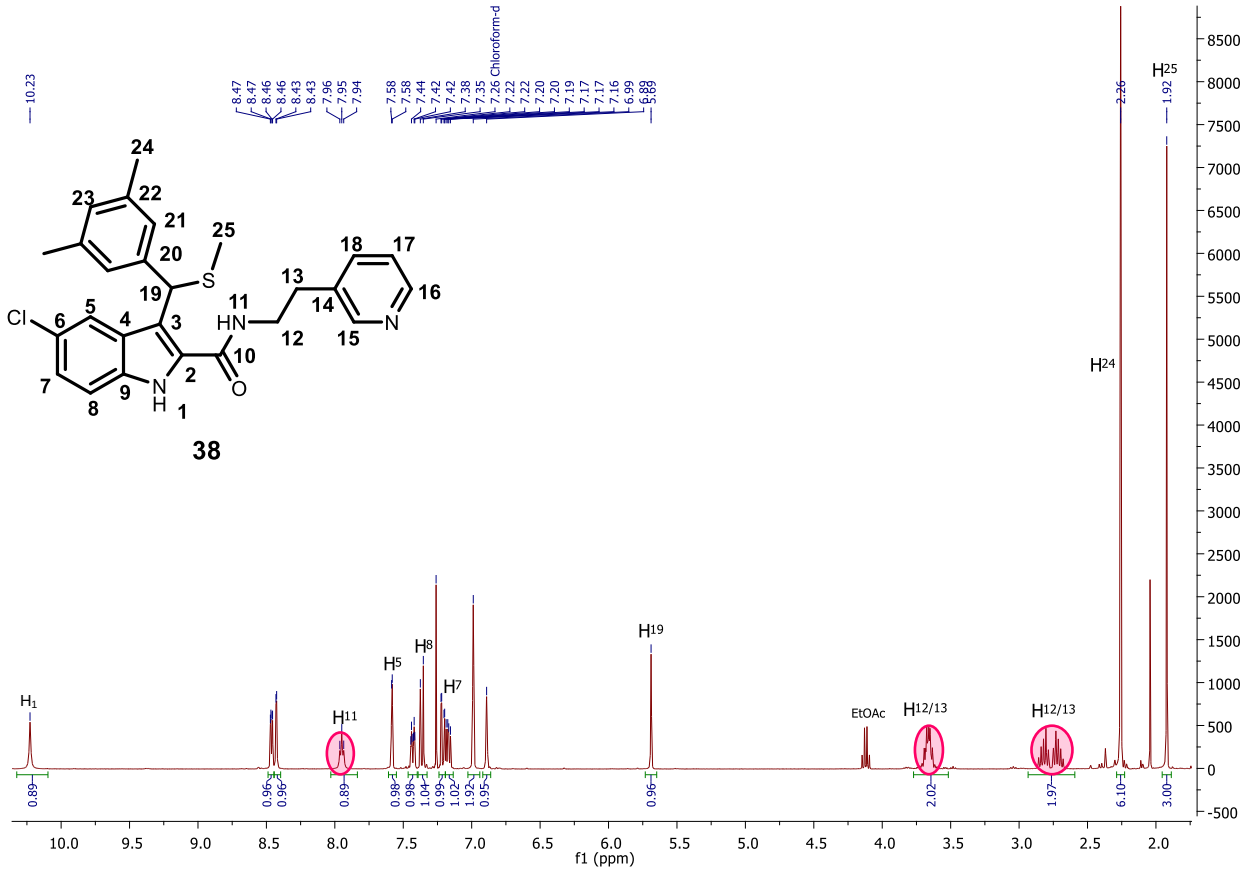


*Figure 2.17: The ^{1}H NMR spectrum of product **38** shown from 10.3 ppm to 1.8 ppm. From left to right the signals correspond to H_1, ArH, H_{11}, H_5, ArH, H_8, H_7, ArH, ArH, ArH, H_{19}, $H_{12/13}$, $H_{12/13}$, H_{24} and H_{25} as labelled. Key proton signals H_{11}, H_{12} and H_{13} are circled in pink.*

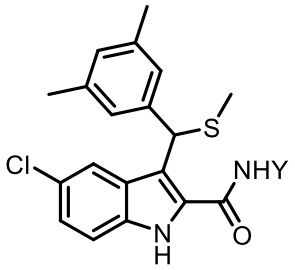
2.3.8 Biological results and concluding remarks

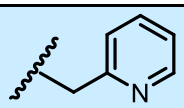
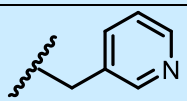
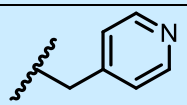
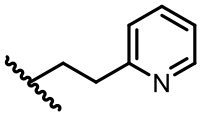
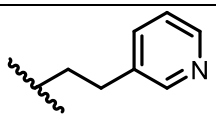
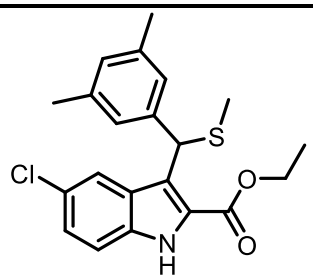
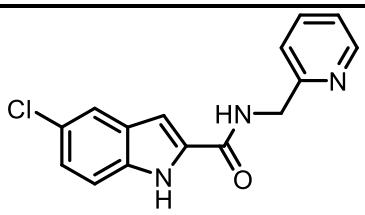
With the synthesis of the five target thioether-containing indole compounds and one "test reaction" compound completed, biological evaluation was carried out by our collaborator, Dr. Adriaan Basson, as introduced in Section 2.2.6. Again, an *in vitro* single-cycle, non-replicative phenotypic assay, using a HIV-1 retroviral vector system for the production of virus-like particles, was utilised.[1] The compounds were incubated with 293T cells together with the virus-like particles. After 48 hours, inhibition was quantified by luminescence measurement. The results obtained are summarised in *Table 2.5* below.

The five target compounds proved to be active against wild-type HIV-1, with compounds **34** – **36** (highlighted in blue, *Table 2.5*) being the most potent. A significant decrease in activity (6 x) was observed for compounds **37** and **38**, with an ethylene bridge instead of the methylene bridge between the amide and pyridinyl functional groups present in the **Y** position of compounds **34** – **38**. This may indicate that the NNRTI-BP is unable to comfortably accommodate this longer chain or that the pyridinyl nitrogen is no longer in the correct position to interact with a residue important for binding. Additionally, with an increase in chain length, the number of rotatable bonds increase requiring a larger reduction in entropy upon binding. This in turn may be associated with a poorer

binding affinity. The importance of the (3,5-dimethylphenyl)(methylthio)methyl moiety is highlighted by the toxicity, and thus inactivity, of compound **56**, lacking this group.

Table 2.4: A summary of the toxicity and activity data obtained from biological testing against wild-type HIV-1 for the final compounds synthesised in Section 2.3.



Compounds		Toxicity (CC_{50}, μM)				Activity (IC_{50}, μM)			
#	Y	1	2	Ave	SD	1	2	Ave	SD
34		23.17	22.51	22.84	0.47	0.044	0.043	0.043	0.001
35		19.13	18.30	18.71	0.59	0.028	0.031	0.030	0.002
36		20.36	18.76	19.56	1.13	0.030	0.037	0.034	0.006
37		21.45	20.49	21.19	0.36	0.259	0.264	0.262	0.003
38		20.71	17.21	18.96	2.47	0.308	0.194	0.251	0.081
NVP ctrl.						0.105	0.076	0.091	0.020
Reference compound **15** and test compound **56**									
15				32.00				0.060	
56		>100	>100	>100	-	13.385	8.008	10.696	3.802

Ave = average

SD = standard deviation

NVP ctrl. = Nevirapine control (IC_{50} value within expected range)

Blue highlight = compound most potent against wild-type HIV-1.

The three compounds **34** – **36**, which proved to be highly active against wild-type HIV-1 and superior to the indole sulphide lead compound **15** (0.060 μM), were next evaluated against a panel of mutant HIV-1 strains harbouring K103N, Y181C, V106M, Y188C, Y188H and G190A single or double amino acid mutations in the RT. These results are summarised in *Table 2.6* below and are visually represented as fold change values for compounds **34** – **36** in *Chart 2.1*. Compounds **35** and **36** inhibited both the mutant V106M and G190A HIV-1 strains, which cause resistance to efavirenz and nevirapine respectively, while resistance was displayed towards compound **34**. Compound **36** additionally remained potent against the Y181C mutation, which renders nevirapine ineffective. All three compounds inhibited the mutant Y188C/H HIV-1 strains. The most exciting results were obtained when the three compounds were evaluated against the particularly problematic and most prevalent mutation for NNRTIs, K103N. Both compounds **35** and **36** remained effective, while compound **34** was only slightly less potent against this mutant strain. Across all mutant strains, except for the double mutant K103N+Y181C, compounds **35** and **36** were more potent than compound **34**, as was the case against wild-type HIV-1.

*Table 2.5: A summary of the activity data obtained from biological testing against a panel of mutant HIV viral strains for the final compounds **34** - **36** synthesised in Section 2.3.*

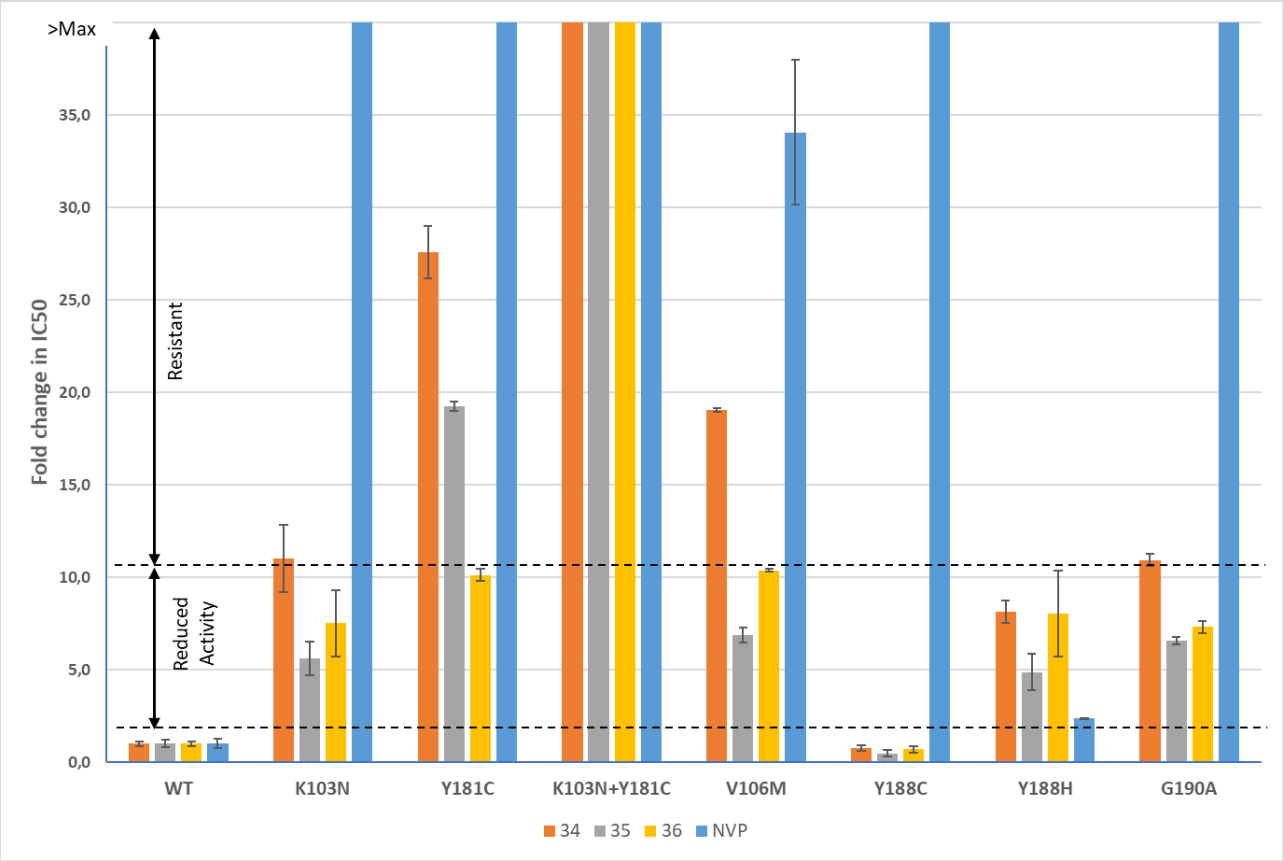
Compounds		Activity (IC_{50}, μM)							
		WT	K103N	Y181C	K103N+Y181C	V106M	Y188C	Y188H	G190A
34	Ave	0.047	0.517	1.294	4.029	0.893	0.036	0.382	0.513
	SD	0.006	0.086	0.067	0.061	0.005	0.007	0.028	0.015
	FC	1.0	11.0	27.6	85.8	19.0	0.8	8.1	10.9
35	Ave	0.031	0.176	0.604	2.798	0.215	0.019	0.152	0.206
	SD	0.006	0.029	0.008	0.214	0.013	0.004	0.031	0.006
	FC	1.0	5.6	19.3	89.3	6.9	0.6	4.9	6.6
36	Ave	0.031	0.235	0.317	2.284	0.325	0.022	0.252	0.229
	SD	0.004	0.056	0.011	0.272	0.002	0.006	0.073	0.011
	FC	1.0	7.5	10.1	72.9	10.4	0.7	8.0	7.3

Ave = average

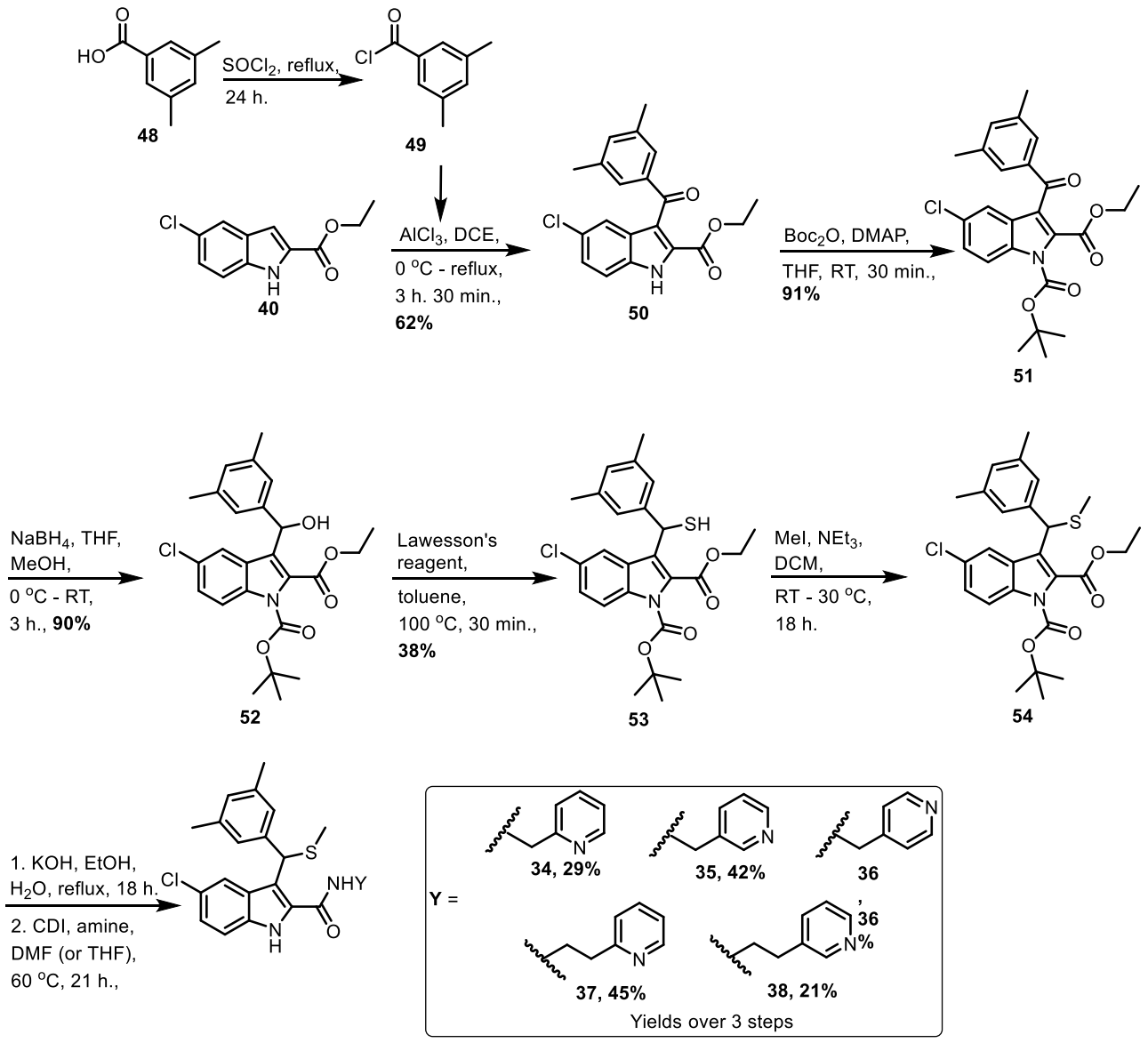
SD = standard deviation

FC = Fold change in IC_{50} relative to the wild-type

*Chart 2.1: Potency expressed as fold change values for compounds **34** – **36**, evaluated against a panel of mutant HIV viral strains. A fold change of 10 or higher for a particular compound indicates viral resistance to the compound.*[42,43]



From our efforts to improve the activity of the indole sulfides, previously synthesised in our group as potent NNRTIs,[10] the indole 2-position was successfully modified with various pyridinyl groups, as summarised in *Scheme 2.14*. After the full characterisation of these compounds, biological evaluation was carried out against wild-type HIV-1. The newly introduced pyridinyl groups improved the activity of the indole sulphide lead compound in three of the five cases (**34** – **36**). With an ester in the indole 2-position (**15**), an activity of 0.060 μM was achieved, as opposed to the best pyridinyl derivative (**35**) with the activity of 0.030 μM. Information about the positioning of the pyridinyl nitrogen and the optimum chain length for the indole 2-position was gathered and is likely to be valuable for future studies. Furthermore, tests against a panel of mutant HIV viral strains revealed that compounds **35** and **36** inhibited the mutant V106M, Y188C/H, G190A and K103N HIV-1 strains of which the results regarding K103N was especially promising, considering its high incidence for NNRTIs in clinical use.



*Scheme 2.14: The complete synthesis of the five target indole sulfide derivatives with pyridinyl moieties as **Y**.*

2.4 Experimental section

The supplementary information pertaining to Sections 2.2 and 2.3 follows. It should be noted that, where applicable, the compounds were obtained as racemic mixtures and that isomers and rotamers complicated the spectra of some of the compounds. Also, an arbitrary numbering system was used and reference to the provided structures should be made when studying the NMR spectra.

2.4.1 General procedures

2.4.1.1 The purification of reagents and solvents

All chemicals used to complete the synthesis reported were purchased from Merck or Sigma Aldrich. Solvents used for chromatographic purposes or work-up procedures were distilled before use by means of conventional distillation procedures. Unless otherwise stated, solvents used for reaction purposes were dried over an appropriate drying agent and then distilled under nitrogen gas.

Tetrahydrofuran was distilled from sodium wire using benzophenone as an indicator. Dichloromethane, 1,2-dichloroethane and dimethylformamide were distilled from calcium hydride. Ethanol was distilled from magnesium turnings and iodine. Diethyl ether and pyridine were bought with a ≥98% purity grade from Sigma Aldrich and were dried on activated 3 Å molecular sieves. *n*-Butyllithium (*n*-BuLi) was bought from Sigma Aldrich with a purity grade of 2.5 M in hexane. Before every use, the *n*-BuLi was titrated three times using either benzylbenzamide or menthol and bipyridine and an average concentration was calculated.[44]

2.4.1.2 Chromatography

Thin layer chromatography (TLC) was performed using Merck silica gel 60 F254 coated on aluminium sheets. Compounds on the TLC plates were viewed under UV light and if necessary, *p*-anisaldehyde, potassium permanganate, or ninhydrin stains were used to visualise the compounds.

Separation of compounds by column chromatography was performed using Merck silica gel (particle size 0.063-0.200 mm, 60 Å) and combinations of hexane and ethyl acetate as solvent.

All compounds submitted for efficacy evaluation (HIV phenotypic assay) were subjected to purity determination by HPLC to ensure a minimum purity of 90%. Purity determination made use of a Waters Acquity UPLC fitted with a photodiode array detector, which was in turn linked to a Waters Synapt G2 QTOF, operated in ESI positive or negative mode with a cone voltage of 15 v. Waters Acquity PDA was set at 280 nm to measure peak purity. A Waters BEH C18 2.1 x 100 mm column was used at 45 °C. The samples were dissolved in the appropriate solvent at a concentration of 1 mg/mL. They were then transferred to a glass vial and 1 μL injected at time 0.
Solvent A: 0.1% formic acid; Solvent B: Acetonitrile with 0.1% formic acid.

2.4.1.3 Spectroscopic and physical data

Nuclear magnetic resonance spectra (^{1}H, ^{13}C) were recorded on a 300 MHz Varian VNMRS (75 MHz for ^{13}C) or a 400 MHz Varian Unity Inova (101 MHz for ^{13}C). Chemical shifts (δ) are reported in ppm and *J*-values are given in Hz. Chemical shifts were recorded using the residual solvent peak or external reference. All spectra were obtained at 25 °C and spectroscopic data were processed using MestReNova v6.0.2-5475. Mass spectrometry was performed on a Waters SYNAPT G2. Infrared spectra were recorded on a Thermo Nicolet Nexus 470 by means of Attenuated Total Reflectance (ATR) mode. Melting points were obtained using a Gallenkamp Melting Point Apparatus.

2.4.1.4 Molecular modelling

Molecular modelling was carried out using the Schrodinger small molecule drug discovery suite, as launched from the Maestro interface. The receptors utilized were obtained from the PDB (2WON and 2RF2) and subjected to the protein preparation protocol. Docking of the various ligands was carried out using Glide XP. Binding energy calculation were carried out using the MMGBSA protocol.

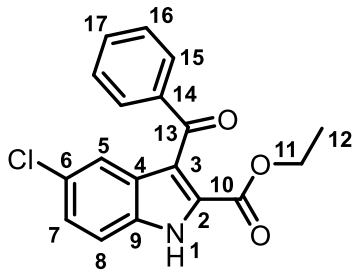
2.4.1.5 Other general procedures

All reactions were carried out under a positive pressure of nitrogen gas, unless water was used as a solvent. The glassware was oven-dried (80 °C) overnight before being purged with N_2 gas. Solvents were removed using a rotary evaporator, followed by removal of trace amounts of solvent using a high vacuum pump at *ca*. 0.08 mm Hg.

2.4.2 Experimental section pertaining to section 2.2

2.4.2.1 Synthesis of ethyl 3-benzoyl-5-chloro-1*H*-indole-2-carboxylate (41)

To a 50 mL three-neck round-bottom flask fitted with a stopper, septum and condenser, which had been flushed with alternating vacuum and nitrogen purge cycles, was added dry 1,2-dichloroethane (DCE) (10 mL). The solution was then cooled to 0 °C using an ice-bath. Anhydrous aluminium chloride ($AlCl_3$) (3 eq., 0.894 g, 6.71 mmol) was added, followed by the dropwise addition of benzoyl chloride (3 eq., 0.78 mL, 6.9 mmol) by means of a syringe through the septum. The ice-bath was then removed and the reaction mixture was left to stir for 30 min. A clear orange reaction mixture resulted. Ethyl 5-chloro-1*H*-indole-2-carboxylate (**40**) (0.500 g, 2.24 mmol) was then added and the reaction mixture was heated at reflux for 4 h. The heat was then turned off and the reaction mixture was left to cool to RT. A saturated solution of $NaHCO_3$ (40 mL) was added to the reaction mixture, which was then filtered through Celite into a 500 mL round-bottom flask. The Celite was washed with ethyl acetate (EtOAc) (3 x 20 mL). The contents of the 500 mL round-bottom flask was then transferred to a separating funnel and the organic and aqueous layers were separated. The aqueous layer was extracted with EtOAc (3 × 30 mL) and the combined organic layers were washed with brine, dried over magnesium sulfate ($MgSO_4$) and filtered. Purification was done by column chromatography (1 - 15% EtOAc/Hexane) to yield the title compound (0.530 g, 1.62 mmol, 72%) (R_f = 0.21, 20% EtOAc/Hexane) as a light yellow solid.



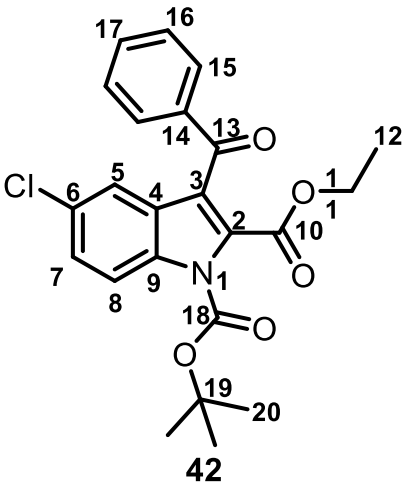
41

Mp 165 – 167 °C. **IR (ATR, cm^{-1})** 3275 (N-H str), 3064 – 2938 (C – H str), 1692 (C=O str), 1642 (C=C str), 1523 (N – H bend), 1450, 1427, 1336, 1256 (C – O str), 1216. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.42 (s, 1H, H_1), 7.89 – 7.84 (m, 2H, 2 x ArH), 7.72 (d, *J* = 2.0 Hz, 1H, H_5), 7.62 – 7.53 (m, 1H, ArH), 7.49 – 7.38 (m, 3H, H_8 and 2 x ArH), 7.34 (dd, *J* = 8.8, 2.0 Hz, 1H, H_7), 4.05 (q, *J* = 7.2 Hz, 2H, H_{11}), 0.87 (t, *J* = 7.2 Hz, 3H, H_{12}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 192.5 (C_{13}), 161.1 (C_{10}), 139.3, 133.9 (C_{14}), 133.2, 129.6, 128.6, 128.4, 128.3, 127.7, 127.0, 121.5 (C_5), 119.4, 113.2 (C_8), 61.9 (C_{11}), 13.5 (C_{12}).

Spectra for this compound matched those reported in the literature.[1,9,10,17]

2.4.2.2 Synthesis of 1-(*tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (42)

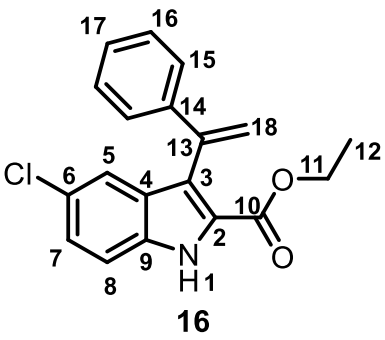
3-Benzoyl-5-chloro-1*H*-indole-2-carboxylate (**41**) (0.518 g, 1.58 mmol) and 4-dimethylaminopyridine (DMAP) (3 crystals, catalytic) were added to a 50 mL two-neck round-bottom flask, which had been flushed with alternating vacuum and nitrogen purge cycles. Dry tetrahydrofuran (THF) (10 mL) was added, followed by di-*tert*-butyl dicarbonate (Boc anhydride) (1.3 eq., 0.47 mL, 2.1 mmol). The reaction mixture was stirred at 30 °C for 30 min., resulting in an orange solution. The solvent was removed *in vacuo* and the crude indole was purified by column chromatography (1 – 5% EtOAc/Hexane) to afford the title compound (0.607 g, 1.42 mmol, 90%) (R_f = 0.48, 20% EtOAc/Hexane) as a clear oil/white semi-solid.



^{1}H NMR (300 MHz, $CDCl_3$) δ 8.11 (dd, *J* = 8.9, 0.4 Hz, 1H, ArH), 7.84 – 7.76 (m, 2H, 2 x ArH), 7.63 (d, *J* = 2.0 Hz, 1H, H_5), 7.59 (ddd [app. dt], *J* = 2.8, 2 x 1.8 Hz, 1H, ArH), 7.53 – 7.43 (m, 2H, 2 x ArH), 7.39 (dd, *J* = 9.0, 2.0 Hz, 1H, H_7), 3.93 (q, *J* = 7.2 Hz, 2H, H_{11}), 1.63 (s, 9H, H_{20}), 1.07 (t, *J* = 7.2 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 191.0 (C_{13}), 161.4 (C_{10}), 148.4 (C_{18}), 138.5, 134.2, 133.8, 133.4, 130.2, 129.3, 128.7, 128.0, 127.3, 121.5, 121.2, 116.4 (C_8), 86.6 (C_{19}), 62.4 (C_{11}), 27.9 (C_{20}), 13.6 (C_{12}). Spectra for this compound matched those reported in the literature.[1,9,10,17]

2.4.2.3 Synthesis of ethyl 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (16)

To dry THF (30 mL) cooled to 0 °C in a 100 mL three-neck round-bottom flask fitted with a rubber septum, which had been flushed with alternating vacuum and nitrogen purge cycles, was added methyltriphenylphosphonium bromide (3.5 eq, 3.71 g, 10.4 mmol). *n*-Butyllithium (*n*-BuLi) (3 eq., 1.88 M, 4.73 mL, 8.89 mmol) was then added dropwise with a syringe. The reaction mixture was heated to 30 °C and was stirred at this temperature for 2 h. The solution containing the ylide turned dark red. 1-(*Tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (**42**) (1.27 g, 2.96 mmol) and THF (30 mL) were added to a separate nitrogen-flushed 100 mL three-neck round-bottom flask, fitted with a dropping funnel. The solution was cooled to 0 °C on ice before the solution containing the ylide was transferred to the dropping funnel with a syringe and was added dropwise. The reaction mixture was heated to 30 °C and was stirred for 2 h. The reaction mixture was quenched with saturated aqueous ammonium chloride (NH_4Cl) solution (40 mL) and the product was extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine, dried over $MgSO_4$ and filtered. After removal of the solvent the crude product was purified by column chromatography (1 – 5% EtOAc/Hexane) to yield the pure title compound (0.703 g, 2.16 mmol, 73%) (R_f = 0.41, 20% EtOAc/Hexane) as a yellow solid.



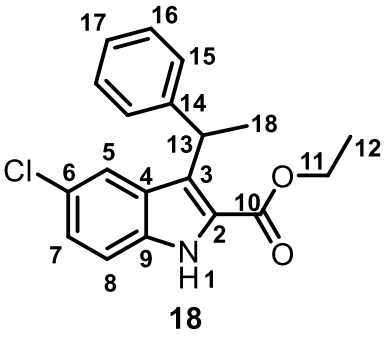
16

^{1}H NMR (300 MHz, $CDCl_3$) δ 9.15 (s, 1H, H_1), 7.61 (d, *J* = 1.9 Hz, 1H, ArH), 7.38 (dd, *J* = 8.8, 0.5 Hz, 1H, H_7), 7.35 – 7.23 (m, 6H, 6 x ArH), 5.99 (d, *J* = 1.3 Hz, 1H, H_{18a}), 5.41 (d, *J* = 1.3 Hz, 1H, H_{18b}), 4.13 (q, *J* = 7.1 Hz, 2H, H_{11}), 1.02 (t, *J* = 7.1 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.8 (C_{10}), 141.5 (C_{13}), 140.9, 134.1, 129.8, 128.3, 127.7, 126.8, 126.5, 126.4, 125.2, 122.8, 121.1, 117.3, 113.0, 61.2 (C_{11}), 13.9 (C_{12}).

Spectra for this compound matched those reported in the literature.[1,9,17]

2.4.2.4 Synthesis of ethyl 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylate (18)

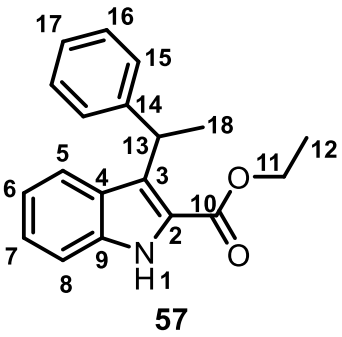
To ethyl 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (**16**) (0.200 g, 0.614 mmol) dissolved in ethanol (EtOH) (10 mL) was added palladium on carbon (Pd/C) (spatula tip) and the system was placed under a hydrogen (H_2) atmosphere (balloon). The reaction mixture was left to stir at RT for 2 h. The reaction mixture was then filtered through Celite. The pure product (0.0733 g, 0.224 mmol, 75%) (R_f = 0.38, 20% EtOAc/Hexane) was obtained as an off-white solid after column chromatography (1 – 5% EtOAc/Hexane).



18

MP 89 – 90 °C. **IR (ATR, cm^{-1})** 3397 (N-H str), 1709 (C=O str), 1538, 1432, 1233 (C-O str). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.81 (s, 1H, H_1), 7.48 – 7.11 (m, 8H, 8 x ArH), 5.36 (q, *J* = 7.3 Hz, 1H, H_{13}), 4.43 (q, *J* = 7.1 Hz, 2H, H_{11}), 1.80 (d, *J* = 7.3 Hz, 3H, H_{18}), 1.41 (t, *J* = 7.1 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.1 (C_{10}), 144.9, 134.6, 128.4, 128.1, 127.6, 127.5, 126.1, 125.9, 125.6, 122.1, 113.0 (C_8), 61.3 (C_{11}), 35.0 (C_{13}), 20.1 (C_{18}), 14.5 (C_{12}). One carbon peak was not visible in the ^{13}C spectrum (it may be C_2, C_3 or C_4). **HRMS** calculated for $C_{19}H_{19}NO_2Cl$ $[M+H]^+$, 328.1104, found 328.1112.

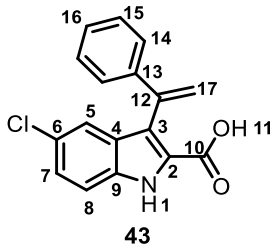
As a by-product, the dechlorinated ethyl 3-(1-phenylethyl)-1*H*-indole-2-carboxylate (**57**) (0.0219 g, 0.0746 mmol, 25%) (R_f = 0.45, 20% EtOAc/Hexane), was produced as a white solid. To keep this product to a minimum, the reaction need to be carefully monitored with TLC.



57

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.75 (s, 1H, H_1), 7.57 – 7.07 (m, 8H, 8 x ArH), 7.01 – 6.91 (m, 1H, ArH), 5.41 (q, *J* = 7.3 Hz, 1H, H_{13}), 4.43 (q, *J* = 7.1 Hz, 2H, H_{11}), 1.83 (d, *J* = 7.3 Hz, 3H, H_{18}), 1.42 (t, *J* = 7.1 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 145.4, 136.4, 128.8, 128.3, 127.6, 126.7, 125.8, 125.4, 123.0, 120.0, 111.9 (C_8), 61.0 (C_{11}), 35.1 (C_{13}), 20.2 (C_{18}), 14.6 (C_{12}). Two carbon peaks were not visible in the ^{13}C spectrum, presumably C_2 or C_3 and C_{10}.

2.4.2.5 Synthesis of 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (43)

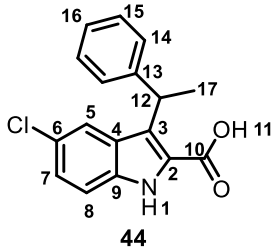


43

To a round bottom flask fitted with a reflux condenser was added ethyl 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (**16**) (0.100 g, 0.307 mmol) and lithium hydroxide (LiOH) (4 eq., 0.0294 g, 1.23 mmol) dissolved in EtOH (2.5 mL) and water (1 mL). The reaction mixture was then heated at reflux for 2 h. The reaction was then diluted with water (10 mL) and the crude product was extracted with EtOAC (3 x 30 mL). The combined organic layers were dried over $MgSO_4$, filtered and the solvent was removed under reduced pressure. The yellow solid (crude product) was used as is in the next reaction.

2.4.2.6 Synthesis of 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylic acid (44)

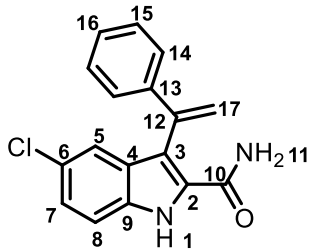
The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**, Section 2.4.2.5).



44

The equivalents used were as follows: ethyl 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylate (**18**) (0.0870 g, 0.265 mmol), LiOH (4 eq., 0.0254 g, 1.06 mmol), EtOH (2.5 mL) and water (1 mL). The light yellow crude product obtained after work-up was used as is in the next reaction.

2.4.2.7 Attempted synthesis of 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (17)



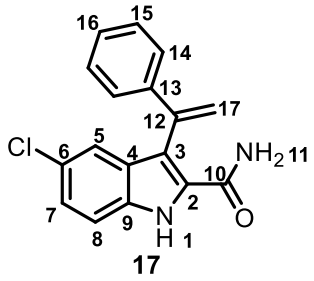
17

Dry chloroform ($CHCl_3$) (5 mL) and dry dimethylformamide (DMF) (0.25 mL, catalytic) was added to a two-neck round-bottom flask fitted with a reflux condenser, after the flask had been flushed with alternating vacuum and nitrogen purge cycles. 5-Chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**) (0.0650 g, 0.218 mmol) and thionyl chloride ($SOCl_2$) (1 eq., 16 mL, 0.22 mmol) were then added and the reaction mixture was left to stir at reflux for 18 h. The reaction mixture was then left to cool to RT, after which it was added dropwise to a 600 mL beaker charged with 25% aqueous ammonium hydroxide solution at 0 °C (on ice). This solution was left to stir on ice for 2 h., after which it was extracted with EtOAc (3 x 15 mL), dried over $MgSO_4$, and filtered. Thin-layer chromatography (TLC) revealed at least seven spots. After column chromatography (30 – 80% EtOAc/Hexane), the fractions collected were not sufficient for analysis by NMR spectroscopy and an alternative synthetic method was investigated.

2.4.2.8 Synthesis of 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (17)

A 20 mL two-neck round-bottom flask fitted with a rubber septum, was flushed with alternating vacuum and nitrogen purge cycles, after which DMF (8 mL) and then 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**) (0.0443 g, 0.149 mmol) were added. Benzotriazol-1-yl-

oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) (1.2 eq., 0.0930 g, 0.179 mmol) and 1-hydroxybenzotriazole hydrate (HOBt·H_2O) (1.2 eq., 0.0241 g, 0.179 mmol) were then added, followed by *N*,*N*-diisopropylethylamine (DIPEA) (1.5 eq., 0.05 mL, 0.2 mmol). Ammonia gas was then bubbled through the reaction mixture for approximately 3 min. and the reaction mixture was left to stir under the ammonia atmosphere for 2 h. at RT. The reaction mixture was diluted with water (5 mL) and extracted with diethyl ether (Et_2O) (3 x 10 mL). The combined organic layers were dried over $MgSO_4$ and filtered and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (50 – 80% EtOAc/Hexane) to yield the pure title compound [0.0460 g, 0.137 mmol, 92% (over 2 steps)] (R_f = 0.62, 100% EtOAc) as an off-white solid.



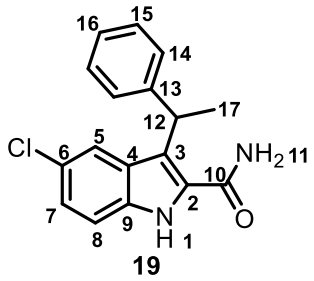
17

MP 156 – 158 °C. **IR (ATR, cm^{-1})** 3437 (C-H str), 3400 - 3100 (N-H str), 1656 (C=O str), 1594, 1573 (N-H bend), 1459. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 10.04 (s, 1H, H_1), 7.47 – 7.28 (m, 7H, 7 x ArH), 7.24 (dd, *J* = 8.8, 2.0 Hz, 1H, H_7), 6.56 (s, 1H, H_{11a}), 6.17 (d, *J* = 1.3 Hz, 1H, H_{17a}), 5.89 (s, 1H, H_{11b}), 5.57 (d, *J* = 1.3 Hz, 1H, H_{17b}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 163.3 (C_{10}), 141.5, 138.8, 133.8, 129.4, 129.0, 127.7, 126.7, 126.5, 125.9, 120.7, 118.9, 118.5, 113.3 (C_8). One carbon was not visible in the ^{13}C spectrum, presumably C_{12}. **HRMS** calculated for $C_{17}H_{14}N_2OCl$ $[M+H]^+$, 297.0795, found 297.0795.

2.4.2.9 Synthesis of 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxamide (19)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (**17**, Section 2.4.2.8).

The equivalents utilised were as follows: 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylic acid (**44**) (0.0200 g, 0.0667 mmol), PyBop (1.2 eq., 0.0420 g, 0.0801 mmol), HOBt·H2O (1.2 eq., 0.0110 g, 0.0801 mmol), DIPEA (1.5 eq., 0.02 mL, 0.1 mmol), DMF (4 mL) and ammonia gas. The pure title product [0.0191 g, 0.0639 mmol, 96% (over 2 steps)] (R_f = 0.57, 100% EtOAc) was obtained as an off-white solid after work-up and purification by column chromatography.



19

MP 114 – 116 °C. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.64 (s, 1H, H_1), 7.64 (d, *J* = 1.6 Hz, 1H, ArH), 7.36 (d, *J* = 8.8 Hz, 1H, ArH), 7.33 – 7.28 (m, 4H, 4 x ArH), 7.25 – 7.20 (m, 2H, 2 x ArH), 5.81 (s, 2H, H_{11}), 4.81 (q, *J* = 7.3 Hz, 1H, H_{12}), 1.86 (d, *J* = 7.3 Hz, 3H, H_{17}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 164.0 (C_{10}), 144.1, 134.0, 129.0, 128.7, 127.6, 127.2, 126.9, 125.7, 125.3, 121.8 (C_7), 120.8 (C_5), 113.2 (C_8), 35.4 (C_{12}), 20.6 (C_{17}). **HRMS** calculated for $C_{17}H_{16}N_2OCl$ $[M+H]^+$, 299.0951, found 299.0951. Note that the amount of material collected was insufficient for IR (ATR, cm^{-1}) analysis.

2.4.2.10 Synthesis of ethyl 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxylate (20)

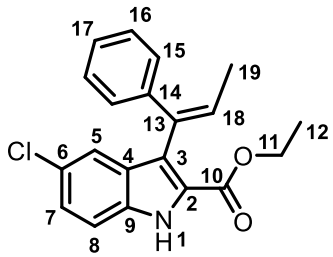
The same procedure was used as for preparing ethyl 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (**16**, Section 2.4.2.3).

The equivalents utilised were as follows:

Ylide formation: Ethyltriphenylphosphonium bromide (3.5 eq., 2.43 g, 6.54 mmol), *n*-BuLi (3 eq., 1.89 M, 3.00 mL, 5.61 mmol) in THF (20 mL).

Wittig reaction: 1-(*tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (**42**) (0.800 g, 1.87 mmol) in THF (20 mL).

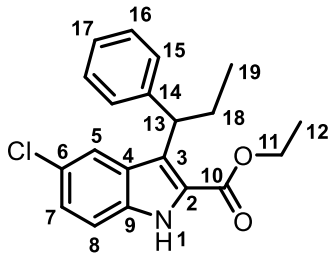
The title compound (0.327 g, 0.961 mmol, 51%) (R_f = 0.43, 20% EtOAc/Hexane) was obtained as an off-white solid. Note that both the *E* and *Z* isomers were isolated as a mixture and that rotamers are also possible. The NMR spectra showed doubling up of numerous peaks and that minor isomer assignments are given in brackets. Due to the complexity of the aromatic region, it was recorded as a large multiplet.



20

MP 119 – 120 °C. **IR (ATR, cm^{-1})** 3308 (N-H str), 3023 – 2853 (C-H str), 1676 (C=O str), 1529 (C=C str), 1459, 1441, 1254 (C-O str), 1194. **^{1}H NMR (400 MHz, $CDCl_3$)** (Major and minor isomers assignments add up to 18 protons) **δ** 9.20 (s, 0.77H, H_1) [9.04 (s, 0.24H, H_1)], 7.60 – 7.15 (m, 8H, 8 x ArH), 6.44 (q, *J* = 7.0 Hz, 0.75H, H_{18}) [5.99 (q, *J* = 7.0 Hz, 0.26H, H_{18})], 4.26 – 4.17 (m, 1.91H, H_{11}) [4.16 – 4.09 (m, 0.095H, H_{11})], 2.01 (d, *J* = 7.0 Hz, 0.60H, H_{19}) [1.63 (d, *J* = 7.0 Hz, 2.41H, H_{19})], 1.20 (t, *J* = 7.1 Hz, 0.68H, H_{12}) [1.12 (t, *J* = 7.1 Hz, 2.37H, H_{12})]. **^{13}C NMR (101 MHz, $CDCl_3$) δ** 161.8 (C_{10}), 142.1, 134.2 [133.8], 133.5 [133.0] (C_{13}), 129.3, 129.0, 128.0, 127.7, 126.8, 126.6, [126.5], 126.2, 126.1, 125.4, 121.1 [120.8] (C_{18}), 113.0 [112.7] (C_8), 61.0 [61.0] (C_{11}), 16.0 [15.5] (C_{12}), [14.2] 14.0 (C_{19}). **HRMS** calculated for $C_{20}H_{19}NO_2Cl$ $[M+H]^+$, 340.1104, found 340.1100.

2.4.2.11 Ethyl 5-chloro-3-(1-phenylpropyl)-1*H*-indole-2-carboxylate (22)

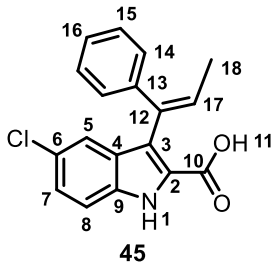


22

Synthesised by Ms Siobhan Brigg during her MSc project.[30]

2.4.2.12 Synthesis of 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxylic acid (45)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**, Section 2.4.2.5).



45

The equivalents utilised were as follows: 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxylate (**20**) (0.0700 g, 0.206 mmol), LiOH (4 eq., 0.0200 g, 0.824 mmol), EtOH (2.5 mL) and water (1 mL).

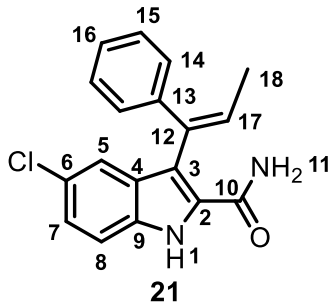
The light yellow crude product obtained after work-up was used as is in the next reaction.

2.4.2.13 Synthesis of 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxamide (21)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (**17**, Section 2.4.2.8).

The equivalents utilised were as follows: ethyl 5-chloro-3-(1-phenylprop-1-en-1-yl)-1*H*-indole-2-carboxylate (**20**) (0.0729 g, 0.234 mmol), PyBop (1.2 eq., 0.146 g, 0.281 mmol), HOBt·H_2O (1.2 eq., 0.0379 g, 0.281 mmol), DIPEA (1.5 eq., 0.06 mL, 0.4 mmol), DMF (5 mL) and ammonia gas. The pure title compound [0.0578 g, 0.186 mmol, 90% (over 2 steps)] (R_f = 0.69, 100% EtOAc) was obtained as a yellow semi-solid after column chromatography.

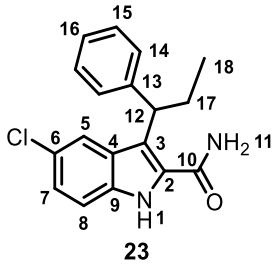
Note that both the *E* and *Z* isomers were isolated as a mixture and that rotamers are also possible. The NMR spectra showed doubling up of numerous peaks and the minor isomer assignments are given in brackets. Due to the complexity of the aromatic region, it was recorded as a large multiplet.



21

IR (ATR, cm^{-1}) 3440, 3256 (N-H str), 1727, 1647 (C=O str), 1581 (C=C str), 1457, 1246. **^{1}H NMR (300 MHz, $CDCl_3$)** (Major and minor isomers assignments add up to 15 protons) **δ** 10.87 (s, 0.80H, H_1) [10.60 (s, 0.20H, H_1)], 7.57 – 7.06 (m, 8H, 8 x ArH), 6.70 (q, *J* = 7.0 Hz, 0.84H, H_{17}), 6.61 (s, 1H, H_{11a}), 6.45 (s, 1H, H_{11b}), [6.14 (q, *J* = 7.0 Hz, 0.16H, H_{17})], [2.05 (d, *J* = 7.0 Hz, 0.65H, H_{18})] 1.69 (d, *J* = 7.0 Hz, 2.35H, H_{18}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** [164.3] 164.1 (C_{10}), 139.7 [138.4], 134.5 [134.0], [134.0] 133.8, [131.0] 129.7, [129.5] 129.2, 128.8 [128.8], 128.6, 127.9 [127.8], 127.8 [127.0], [126.3] 126.1, 125.7 [125.5], [121.7], [120,5] 120.5, 116.1 (C_{17}), 113.6 [113.4] (C_8), 16.1 [16.0], (C_{18}). **HRMS** calculated for $C_{18}H_{16}N_2OCl$ $[M+H]^+$, 311.0951, found 311.0952.

2.4.2.14 5-Chloro-3-(1-phenylpropyl)-1*H*-indole-2-carboxamide (23)



23

Synthesised by Ms Siobhan Brigg during her MSc project.[30]

2.4.2.15 Synthesis of ethyl 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylate (24)

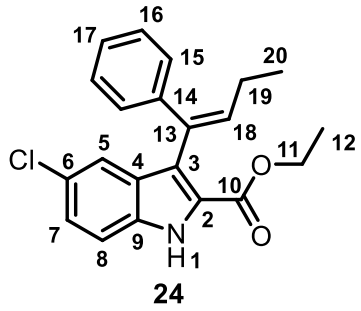
The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylate (**16**, Section 2.4.2.3)

The equivalents utilised were as follows:

Ylide formation: propyltriphenylphosphonium bromide (3.5 eq., 2.52 g, 6.54 mmol), *n*-BuLi (3 eq., 1.89 M, 3.00 mL, 5.61 mmol) in THF, (20 mL).

Wittig reaction: 1-(*tert*-butyl) 2-ethyl 3-benzoyl-5-chloro-1*H*-indole-1,2-dicarboxylate (**42**) (0.800 g, 1.87 mmol) in THF (20 mL).

The title compound (0.274 g, 1.87 mmol, 41%) (R_f = 0.43, 20% EtOAc/Hexane) was obtained as a light yellow solid after purification by column chromatography. Note that both the *E* and *Z* isomers were isolated as a mixture and that rotamers are also possible. The NMR spectra showed doubling up of peaks. Minor isomer given in brackets. Due to the complexity of the aromatic region, it was recorded as a large multiplet.



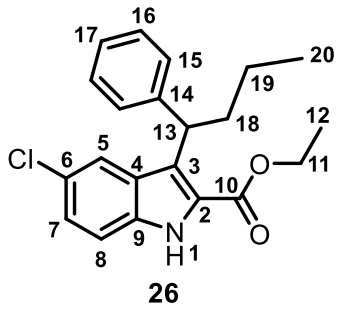
24

MP 106 – 110 °C. **IR (ATR, cm^{-1})** 3315 (N-H str), 2962, 1668 (C=O str), 1529 (C=C str), 1249 (C-O str), 1195, 1099, 1020. **^{1}H NMR (400 MHz, $CDCl_3$)** (Major and minor isomer assignments add up to 20 protons) **δ** 9.16 (s, 0.88H, H_1) [9.02 (s, 0.12H, H_1)], 7.58 – 7.13 (m, 8H, 8 x ArH), 6.35 (t, *J* = 7.5 Hz, 0.89H, H_{18}) [5.86 (t, *J* = 7.5 Hz, 0.11H, H_{18})], 4.28 – 4.17 (m, 1.95H, H_{11}) [4.17 – 4.09 (m, 0.05H, H_{11})], [2.43 (m, 0.19H, H_{19})] 2.07 – 1.82 (m, 1.81H, H_{19}), [1.21 (t, *J* = 7.1 Hz, 0.38H, H_{12})], 1.14 (t, *J* = 7.1 Hz, 2.62H, H_{12}), 0.98 (t, *J* = 7.5 Hz, 3H, H_{20}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.0 (C_{10}), 142.1, [134.3] 134.3, 132.0, 129.4 [129.3], 128.2, 127.8, 126.8 [126.7], [126.4] 126.3, 125.5, 121.2, 121.1, 113.1, 112.9, 110.2, 61.2 (C_{11}), 23.8 (C_{19}), 14.2 (C_{20}), 14.1 (C_{12}). **HRMS** calculated for $C_{21}H_{21}NO_2Cl$ $[M+H]^+$, 354.1261, found 354.1259.

2.4.2.16 Synthesis of ethyl 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylate (26)

The same procedure was used as for preparing ethyl 5-chloro-3-(1-phenylethyl)-1*H*-indole-2-carboxylate (**18**, Section 2.4.2.4).

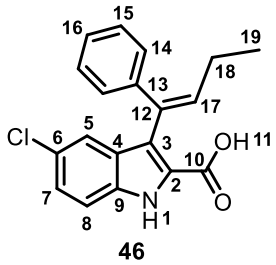
The equivalents utilised were as follows: ethyl 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylate (**24**) (0.112 g, 0.317 mmol), Pd/C (spatula tip), EtOH (5 mL) under H_2 atmosphere (balloon) for 15 min. at RT. The pure product (0.0891 g, 0.250 mmol mmol, 79%) (R_f = 0.45, 20% EtOAc/Hexane) was obtained as an off-white solid after work-up and purification by column chromatography (1 – 5% EtOAc/Hexane).



MP 139 – 141 °C. **IR (ATR, cm^{-1})** 3407 (N-H str), 1712 (C=O str), 1538, 1225 (C-O str). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.97 (s, 1H, H_1), 7.66 (s, 1H, H_5) [note that a small doublet should be seen with higher resolution], 7.43 (d, *J* = 7.6 Hz, 2H, 2 x ArH), 7.36 – 7.11 (m, 5H, 5 x ArH), 5.25 (t, *J* = 7.9 Hz, 1H, H_{13}), 4.46 (q, *J* = 7.1 Hz, 2H, H_{11}), 2.38 – 2.19 (m, 2H, H_{18}), 1.45 (t, *J* = 7.1 Hz, 3H, H_{12}), 1.40 – 1.20 (m, 2H H_{19}), 0.95 (t, *J* = 7.3 Hz, 3H, H_{20}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.4 (C_{10}), 144.6, 134.7, 128.4, 127.9, 127.7, 126.5, 126.1, 125.9, 125.7, 124.8, 122.1, 113.1 (C_8), 61.3 (C_{11}), 40.8 (C_{13}), 36.3 (C_{18}), 21.5 (C_{19}), 14.5 (C_{20}), 14.2 (C_{12}). **HRMS** calculated for $C_{21}H_{23}NO_2Cl$ $[M+H]^+$, 356.1417, found 356.1410.

2.4.2.17 Synthesis of 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylic acid (46)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**, Section 2.4.2.5).

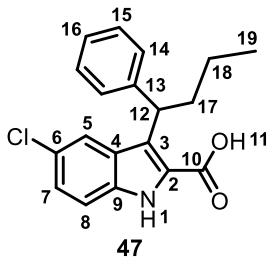


The equivalents used were as follows: ethyl 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylate (**24**) (0.0500 g, 0.141 mmol), LiOH (4 eq., 0.0135 g, 0.564 mmol), EtOH (2.5 mL) and water (1 mL).

The light yellow crude product obtained after work-up was used as is in the next reaction.

2.4.2.18 Synthesis of 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylic acid (47)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxylic acid (**43**, Section 2.4.2.5).



The equivalents used were as follows: ethyl 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylate (**26**) (0.0420 g, 0.118 mmol), LiOH (4 eq., 0.0113 g, 0.472 mmol), EtOH (2.5 mL) and water (1 mL).

The light yellow crude product obtained after work-up was used as is in the next reaction.

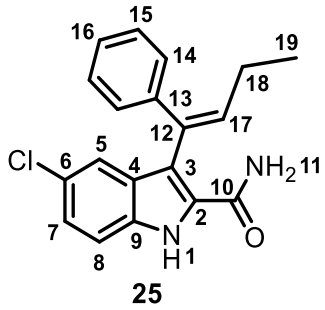
2.4.2.19 Synthesis of 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxamide (25)

The same procedure was utilised as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (**17**, Section 2.4.2.8).

The equivalents used were as follows: 5-chloro-3-(1-phenylbut-1-en-1-yl)-1*H*-indole-2-carboxylic acid (**46**) (0.0518 g, 0.159 mmol), PyBop (1.2 eq., 0.0993 g, 0.191 mmol), HOBt·H_2O (1.2 eq., 0.0258 g, 0.191 mmol), DIPEA (1.5 eq., 0.04 mL, 0.2 mmol), DMF (5 mL) and ammonia gas. The

off-white title product [0.0411 g, 0.127 mmol, 90% (over 2 steps)] (R_f = 0.70, 100% EtOAc) was obtained after work-up and purification by column chromatography.

Note that both the *E* and *Z* isomers were isolated as a mixture and that rotamers are also possible. The NMR spectra showed doubling up of peaks and the minor isomer assignments are given in brackets. Due to the complexity of the aromatic region, it was recorded as a large multiplet.



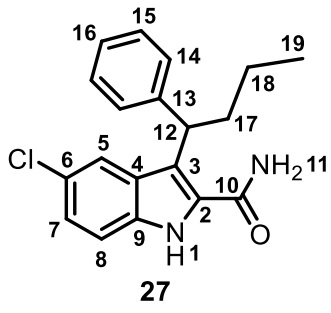
25

MP 80 – 82 °C. **IR (ATR, cm^{-1})** 3441, 3430 – 3100 (N-H str), 1645, 1581 (C=O str), 1455, 1254, 1049, 950. **^{1}H NMR (400 MHz, $CDCl_3$)** (Major and minor isomer assignments add up to a total of 17 protons) **δ** [10.82 (s, 0.02H, H_1)], 10.66 (s, 0.88H, H_1), [10.39 (s, 0.09H, H_1)], 7.53 – 7.13 (m, 8H, 8 x ArH), [6.82 (s, 0.1H, H_{11a})], 6.63 (t, *J* = 7.4 Hz, 0.91H, H_{17}), 6.60 (s, 0.9H, H_{11a}), [6.41 (s, 0.16H, H_{11b})], 6.27 (s, 0.83H, H_{11b}), [6.07 – 5.97 (m, 0.09H, H_{17})], [2.49 (m, 0.09H, H_{18})], 2.07 – 1.95 (m, 1.91H, H_{18}), [1.13 (t, *J* = 7.5 Hz, 0.45H, H_{19})], 0.99 (t, *J* = 7.5 Hz, 2.55H, H_{19}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 163.9 (C_{10}), 139.5, 137.0, 134.2, 132.1 (C_{12}), [129.1], 129.0, 128.9, [128.6], 127.9, 127.7, 126.3, 126.2, 125.8, 120.4, 116.4, 113.6 (C_8), [113.3], [110.1], [95.3], 23.8 (C_{18}), 14.0 (C_{19}). **HRMS** calculated for $C_{19}H_{18}N_2OCl$ $[M+H]^+$, 325.1108, found 325.1102.

2.4.2.20 Synthesis of 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxamide (27)

The same procedure was used as for preparing 5-chloro-3-(1-phenylvinyl)-1*H*-indole-2-carboxamide (**17**, Section 2.4.2.8).

The equivalents utilised were as follows: 5-chloro-3-(1-phenylbutyl)-1*H*-indole-2-carboxylic acid (**47**) (0.0416 g, 0.127 mmol), PyBop (1.2 eq., 0.0792 g, 0.152 mmol), HOBt·H_2O (1.2 eq., 0.0206 g, 0.152 mmol), DIPEA (1.5 eq., 0.03 mL, 0.2 mmol), DMF (5 mL) and ammonia gas. The title product [0.0350 g, 0.107 mmol, 91% (over 2 steps)] (R_f = 0.61, 100% EtOAc) was obtained as an off-white semi-solid after work-up and purification by column chromatography.

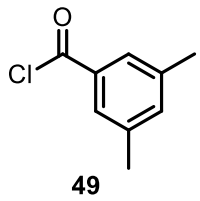


27

IR (ATR, cm^{-1}) 3600 – 3050 (N-H str), 1648, 1585 (C=O str), 1446, 1398, 1337, 1260 (C-O str), 1096, 1057, 946. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.59 (s, 1H, H_1), 7.74 (d, *J* = 1.7 Hz, 1H, H_5), 7.41 – 7.17 (m, 7H, 7 x ArH), 5.86 (s, 2H, H_{11}), 4.59 (t, *J* = 7.9 Hz, 1H, H_{12}), 2.35 – 2.23 (m, 2H, H_{17}), 1.31 – 1.23 (m, 2H, H_{18}), 0.92 (t, *J* = 7.3 Hz, 3H, H_{19}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 164.1 (C_{10}), 150.9, 143.9, 134.0, 129.0, 129.0, 128.3, 127.5, 126.8, 125.8, 125.2, 121.0 (C_7), 120.2 (C_5), 113.2 (C_8), 41.9 (C_{12}), 37.1 (C_{17}), 21.6 (C_{18}), 14.0 (C_{19}). **HRMS** calculated for $C_{19}H_{20}N_2OCl$, $[M+H]^+$, 327.1264, found 327.1263.

2.4.3 Experimental section pertaining to section 2.3

2.4.3.1 Synthesis of 3,5-dimethylbenzoyl chloride (49)

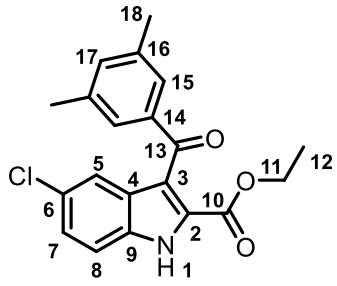


49

3,5-Dimethylbenzoic acid (**48**) (2.00 g, 13.3 mmol) was added to a two-neck 50 mL round-bottom flask containing thionyl chloride (10 mL). The reaction mixture was heated at reflux for 24 h. The solvent was removed using a trolley vacuum pump equipped with a base trap. The crude product was obtained as a brown oil (2.00 g, 11.9 mmol) and was used as is in the next reaction.

2.4.3.2 Synthesis of ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-2-carboxylate (50)

3,5-Dimethylbenzoyl chloride (**49**) (2 eq., 2.00 g, 11.9 mmol) was added to a 100 mL two-neck round-bottom flask fitted with a reflux condenser and charged with DCE (20 mL). The solution was cooled to 0 °C on ice before $AlCl_3$ (2 eq., 1.58 g, 11.9 mmol) was added. The reaction mixture was stirred at 0 °C for 30 min. Ethyl 5-chloro-1*H*-indole-2-carboxylate (1.33 g, 5.93 mmol) was then added and the reaction mixture was heated at reflux for 3 h. The reaction mixture was then cooled to RT and quenched with saturated $NaHCO_3$ solution (30 mL). The emulsion was filtered through Celite into a 500 mL round-bottom flask. The Celite was washed with EtOAc (3 x 20 mL), the contents of the 500 mL round bottom flask was transferred to a separating funnel and the organic and aqueous layers were separated. Extraction was done with EtOAc (3 × 30 mL) and the combined organic layers were washed with brine, dried over $MgSO_4$ and filtered. Purification was done by column chromatography (1 - 15% EtOAc/Hexane) to yield the title compound (1.30 g, 3.65 mmol, 62%) (R_f = 0.26, 20% EtOAc/Hexane) as an off-white solid.



50

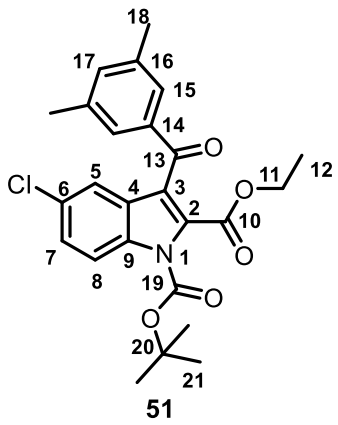
^{1}H NMR (300 MHz, $CDCl_3$) δ 9.66 (s, 1H, H_1), 7.71-7.68 (m, 1H, ArH), 7.49-7.44 (m, 2H, H_{15}), 7.41 (dd, *J* = 8.8, 0.6 Hz, 1H, ArH), 7.31 (dd, *J* = 8.8, 1.9 Hz, 1H, H_7), 7.24 – 7.18 (m, 1H, ArH), 4.09 (q, *J* = 7.2 Hz, 2H, H_{11}), 2.33 (s, 6H, H_{18}), 0.92 (t, *J* = 7.2 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 193.0 (C_{13}), 161.3 (C_{10}), 139.4, 138.2, 134.9, 134.0, 128.1, 127.5, 126.9, 121.4, 119.9, 113.3 (C_8), 61.9 (C_{11}), 21.3 (C_{18}), 13.5 (C_{12}).

Spectra for this compound matched those reported in the literature.[9,10,30]

2.4.3.3 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-1,2-dicarboxylate (51)

Ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-2-carboxylate (**50**) (1.53 g, 4.31 mmol) was dissolved in THF (20 mL) in a 50 mL two-neck round-bottom flask. Boc_2O (1.3 eq., 1.29 mL, 5.60 mmol) was added at RT, followed by DMAP (catalytic). The reaction was monitored by TLC and the starting material was consumed after 30 min. The solvent was removed under reduced pressure

and the product was purified using column chromatography (1 – 15% EtOAc/Hexane) to yield the title compound (1.79 g, 3.94 mmol, 91%) (R_f = 0.51, 20% EtOAc/Hexane) as a yellow solid.



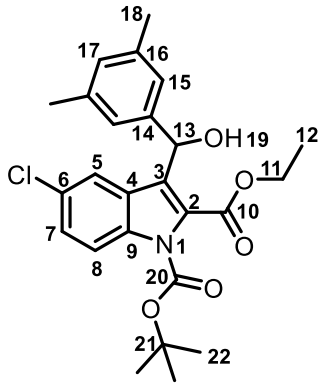
51

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.11 (d, *J* = 9.0 Hz, 1H, ArH), 7.66 (d, *J* = 2.0 Hz, 1H, ArH), 7.41 – 7.37 (m, 3H, 3 x ArH), 7.22 (s, 1H, ArH), 3.91 (q, *J* = 7.2 Hz, 2H, H_{11}), 2.35 (s, 6H, H_{18}), 1.63 (s, 9H, H_{21}), 1.07 (t, *J* = 7.2 Hz, 3H H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 191.4 (C_{13}), 161.4 (C_{10}), 148.5 (C_{19}), 138.6, 138.4, 135.0, 130.2, 128.1, 127.3, 127.1, 121.6, 121.5, 116.4 (C_8), 86.5 (C_{20}), 62.3 (C_{11}), 27.9 (C_{21}), 21.3 (C_{18}), 13.6 (C_{12}).

Spectra for this compound matched those reported in the literature.[9,10,30]

2.4.3.4 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (hydroxy)methyl]-1*H*-indole-1,2-dicarboxylate (52)

1-(*Tert*-butyl) 2-ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-1,2-dicarboxylate (**51**) (1.67 g, 3.67 mmol) was added to a 100 mL round-bottom flask containing THF (5 mL) and MeOH (30 mL). Sodium borohydride ($NaBH_4$) (3 eq., 0.417 g, 11.0 mmol) was added at 0°C and the reaction mixture was warmed to RT. The reaction mixture was quenched with saturated aqueous NH_4Cl solution (30 mL) after 3 h. and the product was extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$, filtered and concentrated. The crude product was purified by column chromatography (1 – 15% EtOAc/Hexane) to yield the title compound (1.52 g, 3.32 mmol, 90%) (R_f = 0.43, 20% EtOAc/Hexane) as an off-white semi-solid.



52

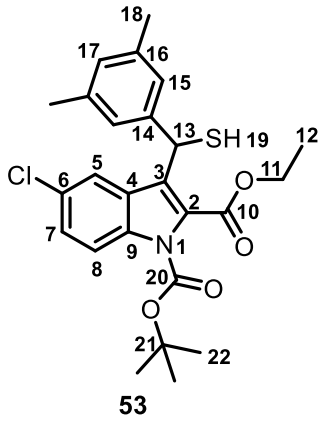
^{1}H NMR (300 MHz, DMSO-d) δ 7.96 (d, *J* = 8.9 Hz, 1H, H_8), 7.73 (d, *J* = 1.3 Hz, 1H, H_5), 7.42 (dd, *J* = 8.9, 1.3 Hz, 1H, H_7), 7.08 (s, 2H, H_{15}), 6.84 (s, 1H, H_{17}), 6.16 (d, *J* = 3.2 Hz, 1H, $H_{13/19}$), 5.99 (d, *J* = 3.2 Hz, 1H, $H_{19/13}$), 4.53 – 4.20 (m, 2H, H_{11}), 2.22 (s, 6H, H_{18}), 1.58 (s, 9H, H_{22}), 1.32 (t, *J* = 7.0 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, DMSO-d) δ** 161.6 (C_{10}), 148.4 (C_{20}), 143.1, 137.1, 134.1, 128.4, 127.5, 127.2, 127.2, 127.0, 126.4, 123.4, 121.8, 116.1 (C_8), 85.5 (C_{21}), 66.8 (C_{13}), 61.7 (C_{11}), 27.4 (C_{22}), 21.1 (C_{18}), 13.9 (C_{12}).

Spectra for this compound matched those reported in the literature.[9,10,30]

2.4.3.5 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (mercapto)methyl]-1*H*-indole-1,2-dicarboxylate (53)

1-(*Tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(hydroxy)methyl]-1*H*-indole-1,2-dicarboxylate (**52**) (1.49 g, 3.25 mmol) was dissolved in toluene (20 mL) in a two-neck round-bottom flask fitted with a reflux condenser. Lawesson's reagent (0.6 eq., 0.788 g, 1.95 mmol) was added and the reaction mixture was heated to 100 °C for 30 min. The reaction mixture was then cooled to RT and the solvent was removed under reduced pressure. The crude product was purified by column

chromatography (1 – 15% EtOAc/Hexane) to yield the title product (0.584 g, 1.23 mmol, 38%) (R_f = 0.54, 20% EtOAc/Hexane) as a yellow oil.



53

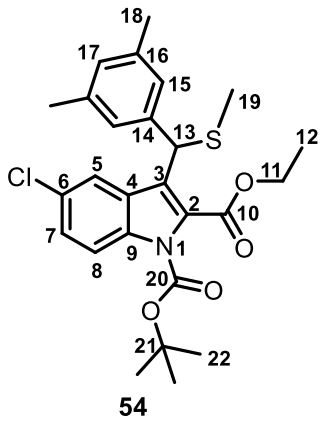
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.02 (dd, *J* = 9.0, 0.4 Hz, 1H, H_8), 7.59 (dd, *J* = 2.1, 0.4 Hz, 1H, H_5), 7.31 (dd, *J* = 9.0, 2.1 Hz, 1H, H_7), 7.14 – 7.09 (m, 2H, H_{15}), 6.90 – 6.86 (m, 1H, H_{17}), 5.81 (d, *J* = 5.6 Hz, 1H, H_{13}), 4.38 (q, *J* = 7.2 Hz, 2H, H_{11}), 2.37 (d, *J* = 5.6 Hz, 1H, H_{19}), 2.28 (s, 6H, H_{18}), 1.64 (s, 9H, H_{22}), 1.38 (t, *J* = 7.2 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.4 (C_{10}), 149.0 (C_{20}), 139.8, 138.2, 135.0, 129.2, 128.7, 127.9, 127.8, 126.8, 125.9, 125.4, 121.8, 116.5, 85.5 (C_{21}), 62.1 (C_{11}), 37.8 (C_{13}), 28.1 (C_{22}), 21.5 (C_{18}), 14.2 (C_{12}).

Spectra for this compound matched those reported in the literature.[10,30]

2.4.3.6 Synthesis of 1-(*tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl) (methylthio)methyl]-1*H*-indole-1,2-dicarboxylate (54)

1-(*Tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(mercapto)methyl]-1*H*-indole-1,2-dicarboxylate (**53**) (0.567 g, 1.20 mmol) was dissolved in dichloromethane (20 mL) in a 50 mL two-neck round-bottom flask. Triethylamine (NEt_3) (2.5 eq., 0.42 mL, 3.0 mmol) was added to the flask and the reaction mixture was stirred for 15 min. before iodomethane (MeI) (2.5 eq., 0.19 mL, 3.0 mmol) was added. The reaction mixture was then heated to 30 °C and was stirred at this temperature for 18 h. The reaction mixture was quenched with saturated aqueous NH_4Cl solution (50 mL) and the crude product was extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (1 – 10% EtOAc/Hexane) to yield the title compound (0.454 g, 0.930 mmol) (R_f = 0.54, 20% EtOAc/Hexane) as a clear oil.

Separating the product from the starting material was difficult since the product and starting material had the same R_f value. The reaction was thus stirred overnight for optimum conversion. Some starting material was, however, observed in the NMR spectra of this product. For this reason, the yield is not reported. The major peaks (product peaks) are reported below.

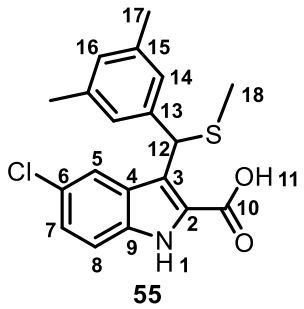


54

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.01 (d, *J* = 9.1 Hz, 1H, H_8), 7.80 (d, *J* = 2.0 Hz, 1H, H_5), 7.30 (dd, *J* = 9.1, 2.0 Hz, 1H, H_7), 7.14 – 7.08 (m, 2H, H_{15}), 6.90 – 6.84 (m, 1H, H_{17}), 5.43 (s, 1H, H_{13}), 4.41 (q, *J* = 7.1 Hz, 2H, H_{11}), 2.27 (s, 6H, H_{18}), 2.02 (s, 3H, H_{19}), 1.64 (s, 9H, H_{22}), 1.38 (t, *J* = 7.1 Hz, 3H, H_{12}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.6 (C_{10}), 149.0 (C_{20}), 138.8, 138.2, 135.0, 129.5, 129.2, 128.6, 128.2, 126.8, 125.8, 125.4, 123.2, 122.3, 116.4, 85.4 (C_{21}), 62.1 (C_{11}), 46.3 (C_{13}), 28.1 (C_{22}), 21.5 (C_{18}), 16.0 (C_{19}), 14.3 (C_{12}).

Spectra for this compound matched those reported in the literature.[10,30]

2.4.3.7 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (55)

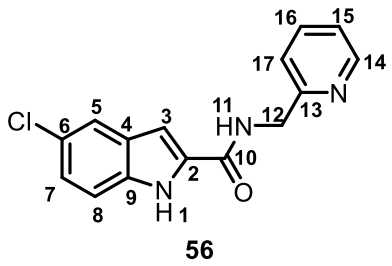


55

1-(*Tert*-butyl) 2-ethyl 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-1,2-dicarboxylate (**54**) (impure) (0.433 g, 0.887 mmol) was added to a round-bottom flask charged with EtOH (20 mL) and water (8 mL). Potassium hydroxide (KOH) (4 eq., 2.00 g, 3.55 mmol) was then added and the flask was fitted with a reflux condenser. The reaction mixture was heated at reflux for 18 h., after which the reaction mixture was cooled to RT. The reaction mixture was then diluted with water (10 mL) and acidified with a 4 M hydrochloric acid (HCl) solution until a pH of 3 was reached. The crude product was extracted with EtOAc (3 x 30 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The title compound was obtained as a red/brown solid and no further purification was done.

2.4.3.8 Synthesis of 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (56)

5-Chloro-1*H*-indole-2-carboxylic acid (kindly supplied by Ms Siobhan Brigg who synthesised this compound during her MSc project[30]) (0.0300 g, 0.153 mmol) was heated in DMF (5 mL) to 60 °C under nitrogen atmosphere in a 25 mL two-neck round-bottom flask. 1,1'-Carbonyldiimidazole (CDI) (1.2 eq., 0.0300 g, 0.184 mmol) was then added and the reaction mixture was stired for 3 h., after which 2-picolylamine (1.2 eq., 0.02 mL, 0.2 mmol) was added and the reaction mixture was left to stir for 18 h. The solvent was then removed by heating under vacuum before the crude product was re-dissolved in EtOAc in order to load the sample on silica. Purification by column chromatography (80 – 100% EtOAc/Hexane) afforded the pure title compound (0.0330 g, 0.115 mmol, 75%) (R_f = 0.44, 100% EtOAc) as an off-white solid.



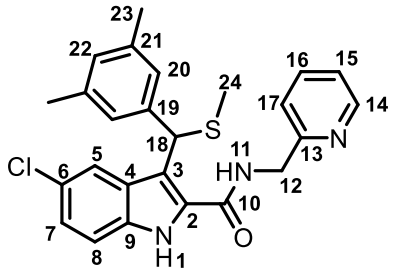
56

MP 189 – 190 °C. **IR (ATR, cm^{-1})** 3265 (N-H str), 1625 (C=O str), 1541, 1413. **^{1}H NMR (300 MHz, DMSO-d) δ** 11.84 (s, 1H, H_1), 9.20 (t, *J* = 6.0 Hz, 1H, H_{11}), 8.57 – 8.48 (m, 1H, H_{pyr}), 7.77 (ddd [app. td], *J* = 2 x 7.7, 1.8 Hz, 1H, H_{pyr}), 7.72 (d, *J* = 2.0 Hz, 1H, H_5), 7.44 (d, *J* = 8.7 Hz, 1H, H_{pyr}), 7.35 (d, *J* = 7.9 Hz, 1H, H_8), 7.31 – 7.24 (m, 1H, H_{pyr}), 7.20 (s, 1H, H_3), 7.17 (dd, *J* = 7.9, 2.0 Hz, 1H, H_7), 4.60 (d, *J* = 6.0 Hz, 2H, H_{12}). **^{13}C NMR (75 MHz, DMSO-d) δ** 160.9 (C_{10}), 158.6, 148.8, 136.8, 134.9, 133.0, 128.1, 124.2, 123.5, 122.2, 121.0, 120.6, 113.8, 102.3, 44.3 (C_{12}). **HRMS** calculated for $C_{15}H_{13}N_3OCl$ $[M+H]^+$, 286.0747, found 286.0740.

2.4.3.9 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (34)

The same procedure was used as for preparing 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**, Section 2.4.3.8).

The equivalents utilised were as follows: 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**) (0.0500 g, 0.139 mmol), CDI (1.2 eq., 0.0270 g, 0.167 mmol), 2-picolylamine (1.2 eq., 0.02 mL, 0.1 mmol), DMF (2 mL). The pure title product [0.0275 g, 0.0611 mmol, 29% (3 steps)] (R_f = 0.52, 100% EtOAc) was obtained as a light yellow solid after work-up and purification by column chromatography (80 – 100% EtOAc/Hexane).



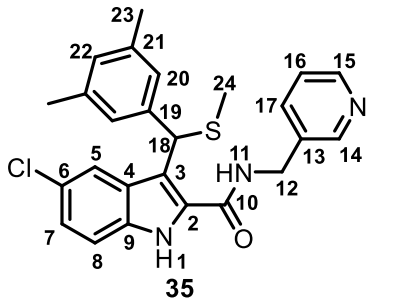
34

MP 152 – 154 °C. **IR (ATR, cm^{-1})** 3450 – 3000 (N-H str), 1556 (C=O str), 1435, 1278, 1257. **^{1}H NMR (300 MHz, CDCl$_3$) δ** 10.58 (s, 1H, H_1), 8.72 – 8.28 (m, 2H, 2 x ArH), 7.67 (s, 1H, ArH), 7.59 (dd [app. t], *J* = 2 x 7.6 Hz, 1H, ArH), 7.35 – 7.00 (m, 6H, 6 x ArH), 6.88 (s, 1H, ArH), 5.94 (s, 1H, H_{18}), 5.04 – 4.60 (m, 2H, H_{12}), 2.24 (s, 6H, H_{23}), 2.05 (s, 3H, H_{24}). **^{13}C NMR (75 MHz, CDCl$_3$) δ** 162.4 (C_{10}), 156.6, 149.3, 138.8, 138.3, 136.9, 133.8, 129.8, 129.3, 129.1, 126.1, 125.9, 125.1, 122.5, 121.7, 120.6, 115.8, 113.4 (C_8), 46.2 (C_{12}), 45.4 (C_{18}), 21.5 (C_{23}), 15.7 (C_{24}). **HRMS** calculated for $C_{25}H_{25}N_3OSCl$ $[M+H]^+$, 450.1407, found 450.1404.

2.4.3.10 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(pyridin-3-ylmethyl)-1*H*-indole-2-carboxamide (35)

The same procedure was used as for preparing 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**, Section 2.4.3.8).

The equivalents utilised were as follows: 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**) (0.0500 g, 0.139 mmol), CDI (1.2 eq., 0.0270 g, 0.167 mmol), 3-picolylamine (1.2 eq., 0.02 mL, 0.2 mmol), THF (2 mL) (note that for practical purposes DMF was substituted for THF). The pure title product [0.0394 g, 0.0876 mmol, 42% (3 steps)] (R_f = 0.29, 100% EtOAc) was obtained as a light yellow solid after work-up and purification by column chromatography (80 – 100% EtOAc/Hexane).



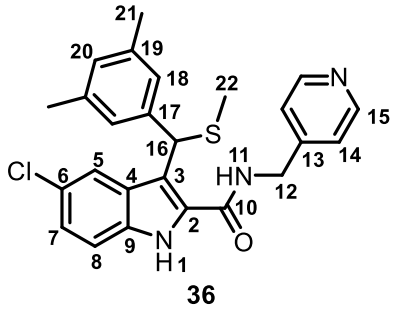
35

MP 60 – 62 °C. **IR (ATR, cm^{-1})** 3500 – 3000 (N-H str), 1633 (C=O str), 1552, 1426, 1255, 1028. **^{1}H NMR (300 MHz, CDCl$_3$) δ** 10.60 (s, 1H, H_1), 8.56 (dd, *J* = 4.6, 1.2 Hz, 1H, ArH), 8.51 (d, *J* = 1.4 Hz, 1H, ArH), 8.40 (t, *J* = 5.7 Hz, 1H, H_{11}), 7.63 (d, *J* = 1.7 Hz, 1H, ArH), 7.40 – 7.29 (m, 2H, 2 x ArH), 7.26 – 7.14 (m, 2H, 2 x ArH), 7.02 (s, 2H, H_{20}), 6.91 (s, 1H, H_{22}), 5.84 (s, 1H, H_{18}), 4.81 – 4.40 (m, 2H, H_{12}), 2.24 (s, 6H, H_{23}), 2.01 (s, 3H, H_{24}). **^{13}C NMR (75 MHz, CDCl$_3$) δ** 162.2 (C_{10}), 149.3, 148.9, 138.5, 138.2, 135.3, 133.4, 133.4, 130.0, 129.6, 129.6, 126.4, 125.7, 125.4, 123.7, 119.5, 114.9, 113.5 (C_8), 45.3 (C_{18}), 41.5 (C_{12}), 21.5 (C_{23}), 15.5 (C_{24}). **HRMS** calculated for $C_{25}H_{25}N_3OSCl$ $[M+H]^+$, 450.1407, found 450.1401.

2.4.3.11 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(pyridin-4-ylmethyl)-1*H*-indole-2-carboxamide (36)

The same procedure was used as for preparing 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**, Section 2.4.3.8).

The equivalents utilised were as follows: 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**) (0.0500 g, 0.139 mmol), CDI (1.2 eq., 0.0270 g, 0.167 mmol), 4-picolylamine (1.2 eq., 0.02 mL, 0.2 mmol), THF (2 mL) (note that for practical purposes DMF was substituted for THF). The pure title product [0.0342 g, 0.0760 mmol, 36% (3 steps)] (R_f = 0.40, 100% EtOAc) was obtained as an off-white solid after work-up and purification by column chromatography (80 – 100% EtOAc/Hexane).



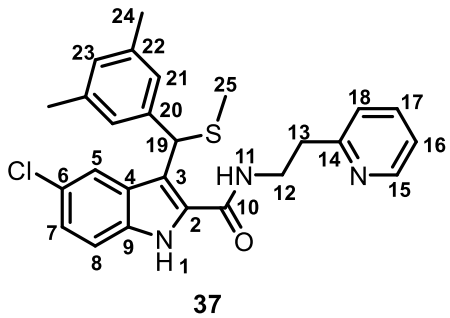
36

MP 163 – 164 °C. **IR (ATR, cm^{-1})** 3232 (N-H str), 3034, 2907, 1640 (C=O str), 1587, 1434, 1313, 1278 (C-O str). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 10.23 (s, 1H, H_1), 8.58 – 8.34 (m, 3H, 3 x ArH), 7.61 (d, *J* = 1.6 Hz, 1H, ArH), 7.33 – 7.17 (m, 2H, 2 x ArH), 7.02 – 6.95 (m, 2H, 2 x ArH), 6.93 – 6.84 (m, 3H, 3 x ArH), 5.83 (s, 1H, H_{16}), 4.81 – 4.30 (m, 2H, H_{12}), 2.21 (s, 6H, H_{21}), 2.00 (s, 3H, H_{22}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.2 (C_{10}), 150.1, 146.9, 138.6, 138.2, 133.2, 130.3, 129.6, 129.5, 126.6, 125.7, 125.6, 122.1, 119.2, 114.8, 113.4 (C_8), 45.1 (C_{16}), 42.7 (C_{12}), 21.5 (C_{21}), 15.4 (C_{22}). **HRMS** calculated for $C_{25}H_{25}N_3OSCl$ $[M+H]^+$, 450.1407, found 450.1420.

2.4.3.12 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(2-(pyridin-2-yl)ethyl)-1*H*-indole-2-carboxamide (37)

The same procedure was used as for preparing 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**, Section 2.4.3.8).

The equivalents utilised were as follows: 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**) (0.0500 g, 0.139 mmol), CDI (1.2 eq., 0.0270 g, 0.167 mmol), 2-(pyridyl)ethylamine (1.2 eq., 0.02 mL, 0.2 mmol), THF (2 mL) (note that for practical purposes DMF was substituted for THF). The pure title product [0.0437 g, 0.0942 mmol, 45% (3 steps)] (R_f = 0.49, 100% EtOAc) was obtained as a light yellow solid after work-up and purification by column chromatography (80 – 100% EtOAc/Hexane).



37

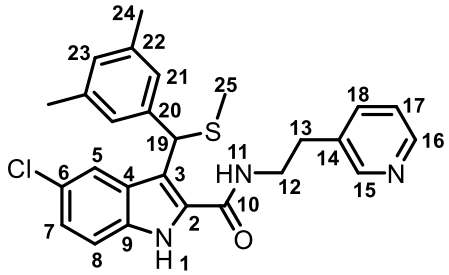
MP 60 – 61 °C. **IR (ATR, cm^{-1})** 3500 – 3000 (N-H str), 1629 (C=) str), 1552, 1434, 1261. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 10.53 (s, 1H, H_1), 8.49 – 8.33 (m, 1H, ArH), 8.07 (t, *J* = 5.4 Hz, 1H, H_{11}), 7.63 (d, *J* = 1.7 Hz, 1H, H_5), 7.55 (ddd [app. td], *J* = 2 x 7.6, 1.8 Hz, 1H, ArH), 7.37 (d, *J* = 8.7 Hz, 1H, H_8), 7.17 (dd, *J* = 8.7, 1.9 Hz, 1H, H_7), 7.13 – 7.05 (m, 2H, 2 x ArH), 7.03 (s, 2H, 2 x ArH), 6.87 (s, 1H, ArH), 5.76 (s, 1H, H_{19}), 3.99 – 3.78 (m, 2H, $H_{12/13}$), 3.16 – 2.90 (m, 2H, $H_{12/13}$), 2.25 (s, 6H, H_{24}), 1.91 (s, 3H, H_{25}). **^{13}C NMR (75 MHz, $CDCl_3$) δ**

162.2 (C_{10}), 159.1, 149.4, 138.7, 138.2, 136.7, 133.8, 130.4, 129.3, 129.1, 126.0, 125.8, 124.9, 123.5, 121.7, 120.6, 114.8, 113.4 (C_8), 46.0 (C_{19}), 39.4 (C_{12}), 37.0 (C_{13}), 21.5 (C_{24}), 15.6 (C_{25}). **HRMS** calculated for $C_{26}H_{27}N_3OSCl$ $[M+H]^+$, 464.1563, found 464.1556.

2.4.3.13 Synthesis of 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-*N*-(2-(pyridin-3-yl)ethyl)-1*H*-indole-2-carboxamide (38)

The same procedure was used as for preparing 5-chloro-*N*-(pyridin-2-ylmethyl)-1*H*-indole-2-carboxamide (**56**, Section 2.4.3.8).

The equivalents utilised were as follows: 5-chloro-3-[(3,5-dimethylphenyl)(methylthio)methyl]-1*H*-indole-2-carboxylic acid (**55**) (0.0500 g, 0.139 mmol), CDI (1.2 eq., 0.0270 g, 0.167 mmol), 3-(2-aminoethyl)pyridine (1.2 eq., 0.02 mL, 0.2 mmol), THF (2 mL) (note that for practical purposes DMF was substituted for THF). The pure title product [0.0206 g, 0.0444 mmol, 21% (3 steps)] (R_f = 0.33, 100% EtOAc) was obtained as a yellow solid after work-up and purification by column chromatography (80 – 100% EtOAc/Hexane).



38

MP 60 – 61 °C. **IR (ATR, cm^{-1})** 3500 – 3100 (N-H str), 1633 (C=O str), 1553, 1425, 1258, 1027, 800. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 10.23 (s, 1H, H_1), 8.46 (dd, *J* = 4.8, 1.5 Hz, 1H, ArH), 8.43 (d, *J* = 2.0 Hz, 1H, ArH), 7.95 (t, *J* = 5.7 Hz, 1H, H_{11}), 7.58 (d, *J* = 1.9 Hz, 1H, H_5), 7.43 (ddd [app. dt], *J* = 7.8, 2 x 1.9 Hz, 1H, ArH), 7.36 (d, *J* = 8.7 Hz, 1H, H_8), 7.21 (dd, *J* = 8.7, 1.9 Hz, 1H, H_7), 7.17 (dd, *J* = 7.7, 4.9 Hz, 1H, ArH), 6.99 (s, 2H, 2 x ArH), 6.89 (s, 1H, ArH), 5.69 (s, 1H, H_{19}), 3.77 – 3.52 (m, 2H, $H_{12/13}$), 2.93 – 2.59 (m, 2H, $H_{12/13}$), 2.26 (s, 6H, H_{24}), 1.92 (s, 3H, H_{25}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 161.9 (C_{10}), 150.1, 148.0, 138.2, 138.1, 136.0, 134.2, 133.1, 129.9, 129.7, 129.3, 126.1, 125.7, 125.1, 123.4, 119.5, 114.3, 113.2, 45.4 (C_{19}), 40.7 (C_{12}), 32.6 (C_{13}), 21.4 (C_{24}), 15.3 (C_{25}). **HRMS** calculated for $C_{26}H_{27}N_3OSCl$ $[M+H]^+$, 464.1563, found 464.1575.

2.5 Bibliography

1 M. Hassam, A. E. Basson, D. C. Liotta, L. Morris, W. A. L. van Otterlo and S. C. Pelly, *ACS Med. Chem. Lett.*, 2012, **3**, 470–475.

2 T. M. Williams, T. M. Ciccarone, S. C. MacTough, C. S. Rooney, S. K. Balani, J. H. Condra, E. A. Emini, M. E. Goldman, W. J. Greenlee, L. R. Kauffman, J. A. O'Brien, V. V. Sardana, W. A. Schleif, A. D. Theoharides and P. S. Anderson, *J. Med. Chem.*, 1993, **36**, 1291–1294.

3 R. Silvestri, G. De Martino, G. La Regina, M. Artico, S. Massa, L. Vargiu, M. Mura, A. G. Loi, T. Marceddu and P. La Colla, *J. Med. Chem.*, 2003, **46**, 2482–2493.

4 R. Silvestri, M. Artico, G. De Martino, G. La Regina, R. Loddo, M. La Colla and P. La Colla, *J. Med. Chem.*, 2004, **47**, 3892–3896.

5 R. Ragno, M. Artico, G. De Martino, G. La Regina, A. Coluccia, A. Di Pasquali and R. Silvestri,

J. Med. Chem., 2005, **48**, 213–223.

6 Z. Zhao, S. E. Wolkenberg, M. Lu, V. Munshi, G. Moyer, M. Feng, A. V. Carella, L. T. Ecto, L. J. Gabryelski, M.-T. Lai, S. G. Prasad, Y. Yan, G. B. McGaughey, M. D. Miller, C. W. Lindsley, G. D. Hartman, J. P. Vacca and T. M. Williams, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 554–559.

7 F.-R. Alexandre, A. Amador, S. Bot, C. Caillet, T. Convard, J. Jakubik, C. Musiu, B. Poddesu, L. Vargiu, M. Liuzzi, A. Roland, M. Seifer, D. Standring, R. Storer and C. B. Dousson, *J. Med. Chem.*, 2011, **54**, 392–395.

8 D. Li, P. Zhan, E. De Clercq and X. Liu, *J. Med. Chem.*, 2012, **55**, 3595–3613.

9 R. Müller, I. Mulani, A. E. Basson, N. Pribut, M. Hassam, L. Morris, W. A. L. van Otterlo and S. C. Pelly, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 4376–4380.

10 S. Brigg, N. Pribut, A. E. Basson, M. Avgenikos, R. Venter, M. A. Blackie, W. A. L. van Otterlo and S. C. Pelly, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 1580–1584.

11 R. Silvestri and M. Artico, *Curr. Pharm. Des.*, 2005, **11**, 3779–3806.

12 G. La Regina, A. Coluccia, A. Brancale, F. Piscitelli, V. Gatti, G. Maga, A. Samuele, C. Pannecouque, D. Schols, J. Balzarini, E. Novellino and R. Silvestri, *J. Med. Chem.*, 2011, **54**, 1587–1598.

13 G. La Regina, A. Coluccia, A. Brancale, F. Piscitelli, V. Famiglini, S. Cosconati, G. Maga, A. Samuele, E. Gonzalez, B. Clotet, D. Schols, J. A. Esté, E. Novellino and R. Silvestri, *J. Med. Chem.*, 2012, **55**, 6634–6638.

14 V. Famiglini, G. La Regina, A. Coluccia, S. Pelliccia, A. Brancale, G. Maga, E. Crespan, R. Badia, B. Clotet, J. A. Esté, R. Cirilli, E. Novellino and R. Silvestri, *Eur. J. Med. Chem.*, 2014, **80**, 101–111.

15 V. Famiglini, G. La Regina, A. Coluccia, S. Pelliccia, A. Brancale, G. Maga, E. Crespan, R. Badia, E. Riveira-Muñoz, J. A. Esté, R. Ferretti, R. Cirilli, C. Zamperini, M. Botta, D. Schols, V. Limongelli, B. Agostino, E. Novellino and R. Silvestri, *J. Med. Chem.*, 2014, **57**, 9945–9957.

16 V. Famiglini, G. La Regina, A. Coluccia, D. Masci, A. Brancale, R. Badia, E. Riveira-Muñoz, J. A. Esté, E. Crespan, A. Brambilla, G. Maga, M. Catalano, C. Limatola, F. R. Formica, R. Cirilli, E. Novellino and R. Silvestri, *J. Med. Chem.*, 2017, **60**, 6528–6547.

17 R. Müller, MSc Dissertation, The Rational Design and Synthesis of Novel HIV Non-nucleoside Reverse Transcriptase Inhibitors, Stellenbosch University, 2013.

18 P. Csomós, L. Fodor, I. Mándity and G. Bernáth, *Tetrahedron*, 2007, **63**, 4983–4989.

19 C. Friedel and J. M. Crafts, *J. Chem. Soc.*, 1877, **32**, 725–791.

20 Z. Huang, L. Jin, H. Han and A. Lei, *Org. Biomol. Chem.*, 2013, **11**, 1810–1814.

21 S. K. Guchhait, M. Kashyap and H. Kamble, *J. Org. Chem.*, 2011, **76**, 4753–4758.

22 M. Rueping and B. J. Nachtsheim, *Beilstein J. Org. Chem.*, 2010, **6**, 1–24.

23 G. Wittig and G. Geissler, *Justus Liebigs Ann. Chem.*, 1953, **580**, 44–57.

24 R. W. Hoffmann, *Angew. Chem. Int. Ed.*, 2001, **40**, 1411–1416.

25 B. E. Maryanoff and A. B. Reitz, *Chem. Rev.*, 1989, **89**, 863–927.

26 P. J. Murphy and J. Brennan, *Chem. Soc. Rev.*, 1988, **17**, 1–30.

27 E. Vedejs and K. A. J. Snoble, *J. Am. Chem. Soc.*, 1973, **95**, 5778–5780.

28 I. Fechete, *Comptes Rendus Chim.*, 2016, **19**, 1374–1381.

29 H. B. Kagan, *Angew. Chem., Int. Ed.*, 2012, **51**, 7376–7382.

30 S. E. Brigg, MSc Dissertation, Lead Optimization of an Indole Based HIV-1 Non-nucleoside Reverse Transcriptase Inhibitor, Stellenbosch University, 2017.

31 E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606–631.

32 P. Zhan, X. Liu, Z. Li, C. Pannecouque and E. De Clercq, *Curr. Med. Chem.*, 2009, **16**, 3903–3917.

33 T. Ozturk, E. Ertas and O. Mert, *Chem. Rev.*, 2007, **107**, 5210–5278.

34 H. C. Brown, *Science*, 1980, **210**, 485–492.

35 H. I. Schlesinger, H. C. Brown, B. Abraham, A. C. Bond, N. Davidson, A. E. Finholt, J. R. Gilbreath, H. Hoekstra, L. Horvitz, E. K. Hyde, J. J. Katz, J. Knight, R. A. Lad, D. L. Mayfield, L. Rapp, D. M. Ritter, A. M. Schwartz, I. Sheft, L. D. Tuck and A. O. Walker, *J. Am. Chem. Soc.*, 1953, **75**, 186–190.

36 Z. Wang, in *Comprehensive Organic Name Reactions and Reagents*, Wiley Online Library, 2010, pp. 1722–1727.

37 T. Nishio, *J. Chem. Soc. Perkin Trans.*, 1993, **1**, 1113–1117.

38 G. A. Sulikowski and M. M. Sulikowski, in *Encyclopedia of Reagents for Organic Synthesis*, eds. P. L. Fuchs, J. W. Bode, A. B. Charette, T. Rovis and L. A. Paquette, Wiley Online Library, 2005, pp. 1–4.

39 C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827–10852.

40 T. Wieland and G. Schmeider, *Ann., 1953, **159**.*

41 R. Paul and G. W. Anderson, *J. Am. Chem. Soc.*, 1960, **82**, 4596–4600.

42 A. E. Basson, S.-Y. Rhee, C. M. Parry, Z. El-Khatib, S. Charalambous, T. De Oliveira, D. Pillay, C. Hoffmann, D. Katzenstein, R. W. Shafer and L. Morris, *Antimicrob. Agents Chemother.*, 2015, **59**, 960–971.

43 Monogram Biosciences, *Phenosense HIV Drug Resistance Assay*, San Francisco, 2014.

44 The University of Texas at El Paso, Lecture Notes, http://utminers.utep.edu, Accessed 18/01/2017.

Chapter 3: The Potential of the Indole Scaffold in the Development of Irreversible Non-nucleoside Reverse Transcriptase Inhibitors

3.1 Our strategy

In this section, the work that enthused our study on irreversible indole-based NNRTIs is discussed and the target compounds are introduced.

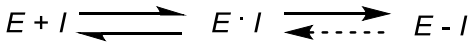
3.1.1 Covalent inhibitors

Irreversible inhibitors, also referred to as covalent inhibitors, are characteristically small molecules containing an electrophilic functional group which is positioned in such a way that nucleophilic reaction with a side chain residue in the NNRTI-BP of the enzyme is possible.[1] Many of the most widely used and highly successful drugs are in fact covalent inhibitors.[1] Aspirin, developed by Bayer in 1897 as an anti-inflammatory agent, is probably the best-known example of such a drug, with the covalent mechanism of action discovered more than 70 years after its commercialisation.[1] Further examples include the penicillins, cephalosporins and fosfomycin, which are antibiotics, as well as omeprazole, a proton pump inhibitor.[1]

A common element in the discovery of numerous covalent drugs is identification by serendipity. Essentially all early covalent drugs were discovered through screening and their covalent mechanisms of action, only revealed later.[2] The number of approved covalent drugs are, however, much lower than that of approved noncovalent drugs, due to concerns for off-target reactivity and toxicity. This was triggered by reports in the 1970s of hepatotoxic properties of some compounds, such as bromobenzene and paracetamol which metabolise to form highly reactive species that covalently bind to liver proteins.[1,2]

Remarkable advantages over noncovalent drugs, such as the lower doses required, fewer side effects and a reduced risk of developing drug resistance, coupled with a clearer understanding of their risks, have resulted in an increase in the number of irreversible drug candidates in clinical trials and in scientific articles.[1] Current covalent drug discovery programmes aim to synthesise targeted covalent inhibitors (TCIs), designed for a specific binding site, as opposed to being discovered by the screening of compound libraries and serendipity.[2] These TCIs require initial reversible binding in the binding pocket (*E* · *I*, *Scheme 3.1*), after which the two reactive centres, now positioned in close proximity with the correct geometry, may react to form a bond resulting in the inhibited complex (*E* – *I*, *Scheme 3.1*).[1,2] Importantly, TCIs bear a bond-forming functional group of low reactivity which may

only bind once the molecule is correctly positioned in the enzyme. This prevents reactions at other sites from occurring. Careful optimization of the non-covalent binding affinity, as well as the reactivity of the electrophilic warhead, is often required.[2]



Scheme 3.1: The mechanism of action of a targeted covalent inhibitor.

Once irreversible inhibition has occurred, enzyme activity will only be restored with the re-synthesis of this enzyme, once all unbound drug has been cleared from the body.[2]

3.1.1.1 Usefulness against the Human Immunodeficiency Virus

As discussed in Section 1.4.2.4.1, Chapter 1, the most common mutations engendering resistance to NNRTIs are the K103N and Y181C mutations. Specifically, the Y181C mutation results in resistance to nevirapine, the first approved NNRTI which is listed as an essential medicine by the World Health Organisation and is used as monotherapy to prevent mother-to-child viral transmission.[3,4] The K103N/Y181C double mutation results in the treatment failure of almost all NNRTIs with the exception of second generation NNRTIs, etravirine and rilpivirine.[4,5] A strategy for combating and eradicating resistance to Y181C and K103N/Y181C bearing mutant strains is therefore of utmost importance. To this end, several drug design groups have attempted to synthesise covalent HIV-1 RT inhibitors that, through modification of Cys181, could inactivate the mutant strains completely.[3,6,7] The next section will describe some strategies that have been tried by researchers to covalently inhibit RT.

MKC-442 (*Figure 3.1*) is an NNRTI that was selected for clinical trials by Triangle Pharmaceuticals, but was withdrawn at phase III, because it did not have superior properties compared to a comparative drug, Abacavir (NRTI), and rapidly selected for mutant viruses, mainly bearing the Y181C mutation.[6,8,9] Analogues with reactive aldehyde groups or epoxides (**58**, **59**, *Figure 3.1*) in the C-5 position were then synthesised by Pedersen *et al*., with the hope of covalent bond formation between the ligand and the Cys181 residue. Unfortunately, no activity against both wild type and mutant strains of the virus was observed.[10] Next, Michael acceptors (**60**, **61**, **62**, *Figure 3.1*) were introduced in the C-6 position. Again, the effort was unsuccessful with no activity observed against the Y181C mutation.[6] Compounds containing an oxime, a reactive aldehyde and a reactive epoxide group in the side chain of the benzylic group (**63**, **64**, **65**, *Figure 3.1*) also did not show activity against the Y181C mutant strain, indicating that covalent binding of the drug to cysteine had not occurred, and, in addition, only moderate activities were observed against the wild-type virus.[7]

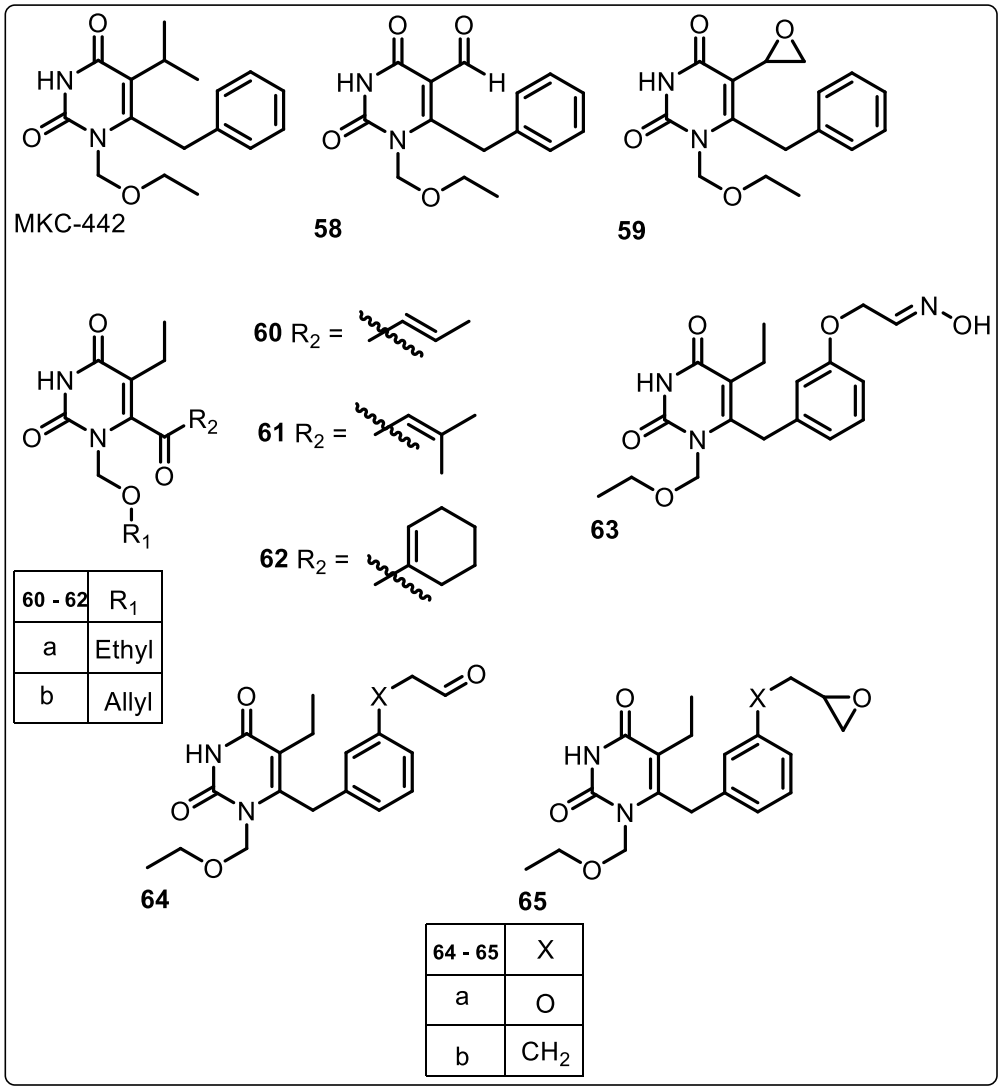


Figure 3.1: MKC-442 analogues synthesised as possible irreversible inhibitors.

Despite these unfruitful results, medicinal chemists did not despair and very recently the departments of Pharmacology, Molecular Biophysics and Biochemistry and Chemistry at Yale University, developed the first irreversible NNRTI.[3] These researchers noticed from the crystal structure of **66** (*Figure 3.2*) in the enzyme that, given the orientation of the molecule, replacement of the naphthyl chloride with an electrophilic warhead could result in covalent modification of Cys181.[3,11] As anticipated, conclusive evidence for covalent bonds between compounds **67** and **68** (*Figure 3.2*) and Cys181 were provided from enzyme inhibition kinetics, mass spectrometry, protein crystallography and antiviral activity in infected human T-cell assays. These compounds are also less toxic than efavirenz and rilpivirine.[3] These positive results are likely to inspire new anti-HIV drugs employing a covalent mechanism of action to destroy the activity of mutant strains bearing the notorious Y181C mutation.

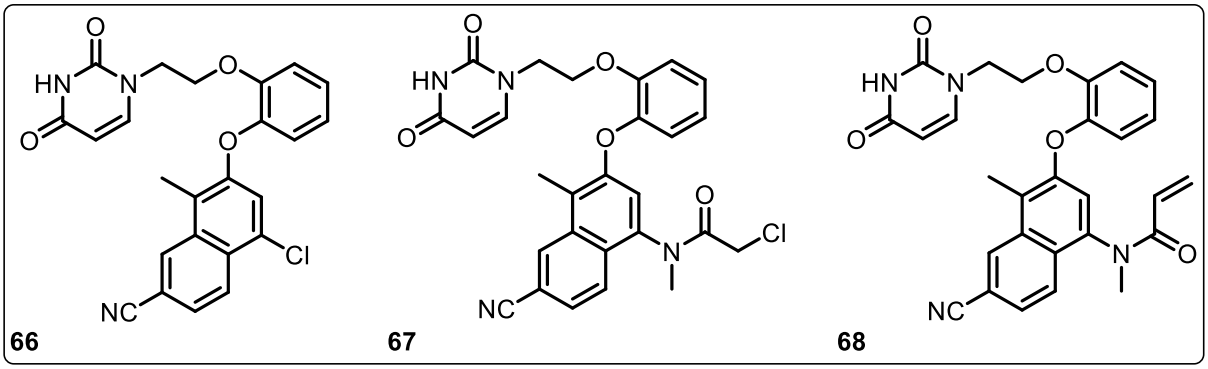


Figure 3.2: The first covalent NNRTIs against the Y181C mutant strain.

3.1.2 Design

With several successful covalent drugs used in long-term therapies for chronic treatment, the use of this concept seemed suitable for HIV-AIDS treatment.[2] A general "recipe" for the development of targeted covalent inhibitors starts with bioinformatic analysis to identify a nucleophilic amino acid, often cysteine, in the binding site. Next, a successful reversible inhibitor, for which the binding mode is known, is then selected as a base-scaffold. Structure-based computational methods are then used to modify the ligand with an electrophilic functional group positioned to react with the identified nucleophilic amino acid. A small set of pilot compounds are then synthesised to test their ability to modify the target enzyme. Importantly, inhibition may thus only be achieved with noncovalent and covalent complementarity in the drug.[2]

To this end, our indole scaffold was proposed as a base-scaffold and electrophilic functional groups were investigated for incorporation at an appropriate site. The most frequently used electrophiles in targeted covalent inhibitors are α,β-unsaturated ketones, vinyl sulfones, α-substituted carbonyl derivatives, bromodihydroisoxazoles, imidazole-1-carboxamides, epoxides, β-lactams and strained lactones.[1] For the synthesis of an initial small set of pilot compounds, α,β-unsaturated amides and an imidazole-1-carboxamide was chosen (*Figure 3.3*), based on the availability of reagents.

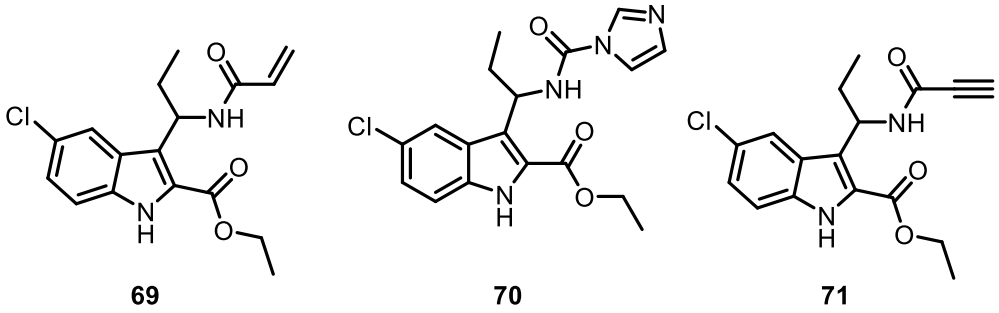
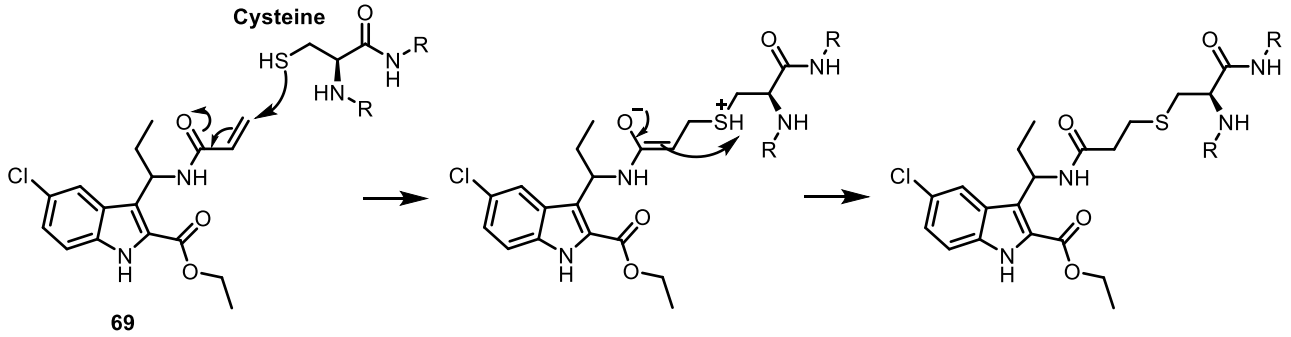


Figure 3.3: A small target library of designed covalent inhibitors as proof-of-concept compounds.

The proposed covalent mechanism of action would involve the attack of the SH group of cysteine (Y181C mutant strain) at the electrophilic functional group incorporated into the indole scaffold.

Covalent modification of a region of the target enzyme was predicted to induce a geometry of the active site unsuitable for catalysis.[1] The covalent bond was proposed to form through a Michael addition reaction with the α,β-unsaturated amides (*Scheme 3.2*) and a simple substitution reaction with imidazole as leaving group in the case of the imidazole-1-carboxamide-containing structure.



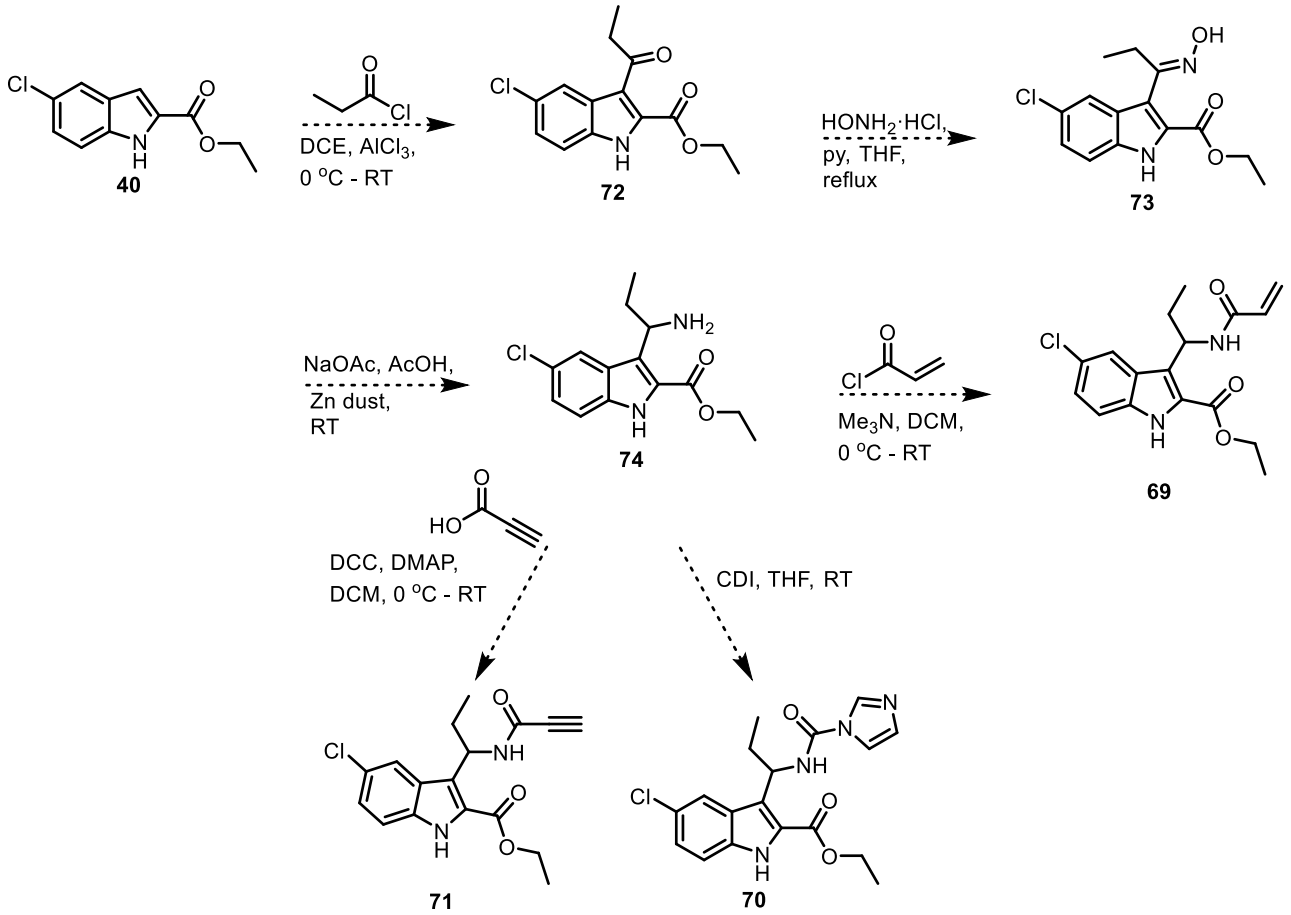
Scheme 3.2: Covalent bond formation between an α, β-unsaturated amide and cysteine 181 in the RT enzyme.

3.2 Targeted covalent inhibitors

The details pertaining to the synthesis of the covalent inhibitors, introduced in Section 3.1.2, are reported in this section. A reaction scheme to introduce the synthetic plan, followed by a detailed discussion of each synthetic step is given. The biological results are then presented and a short summary concludes this section. Note that a supplementary information section is provided at the end of the chapter.

3.2.1 The proposed synthetic route to the target compounds

The well-known and highly effective Friedel-Crafts acylation reaction was proposed for the addition of an ethyl carboxylate to the 3-position of commercially available ethyl 5-chloroindole-2-carboxylate (**40**). Next, an oxime **73** could be formed by the selective reaction of the carbonyl group of the ketone **72** with hydroxylamine. Reduction of the oxime would generate the amine **74**, from which the amides could be formed using acrolyl chloride (**69**), 1,1'-carbonyldiimidazole (CDI) (**70**) and propiolic acid (**71**). This proposed method is outlined in *Scheme 3.3*.

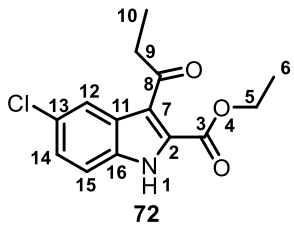


Scheme 3.3: A proposed synthetic strategy towards the target NNRTIs.

3.2.2 Friedel-Crafts acylation reaction

3.2.2.1 Synthesis of ethyl 5-chloro-3-propionyl-1*H*-indole-2-carboxylate (72)[12]

With propionyl chloride and anhydrous aluminium chloride at hand, the reaction with ethyl 5-chloroindole-2-carboxylate (**40**) proceeded without trouble. Propionyl chloride was first dissolved in 1,2-dichloroethane (DCE) and the solution was cooled to 0 °C. Anhydrous aluminium chloride was then added and the reaction mixture was allowed to stir for 30 minutes at this temperature to allow for acylium ion formation. Upon addition of ethyl 5-chloro-1*H*-indole-2-carboxylate (**40**), the ice-water bath was removed and the reaction mixture was stirred at room temperature for 1 hour. General work-up procedures, followed by purification with column chromatography, resulted in the pure title compound. It should be noted that a shorter reaction time was required compared to both the benzoyl and 3,5-dimethylbenzoyl counterparts (**41** section 2.2.2.1 and **50** section 2.3.2.2, Chapter 2, with reaction times of 3 to 4 hours) and the reaction proceeded at room temperature instead of at reflux. A higher yield of 84%, compared to 72% and 62% respectively, was obtained. This could be rationalised by inspection of the acylium ion. A less sterically bulky acylium ion, compared to the benzoyl and 3,5-dimethylbenzoyl counterparts, could result in easier attack by the indole, due to a more accessible electrophile.



72

The product was easily identified from the ^{1}H NMR and ^{13}C NMR spectra. H_9 and H_{10} were observed as a quartet and a triplet integrating for 2 and 3 protons at 3.07 ppm and 1.23 ppm respectively in the ^{1}H NMR spectrum. In addition, C_8 was observed as the furthest downfield peak located at 201.4 ppm in the ^{13}C NMR spectrum. The NMR spectroscopic data collected compared well to those reported in the literature.[12]

3.2.3 Amination

With the goal of introducing an amine into the scaffold, various options came to mind. Two of the most commonly used methods for the preparation of amines are reductive amination of aldehydes or ketones, as is present in **72** (Section 3.2.2.1), and the hydrogenation of nitriles. Both methods go through an imine intermediate. Condensation of a carbonyl compound and an amine forms a hemiaminal before water is eliminated and the imine is formed. Reduction of the imine results in an amine. Similarly, nitriles are hydrogenated first to the imine and then to the amine.[13] In the first case, the imine is commonly reduced using formic acid (the Leuckart-Wallach reaction) or a mild reducing agent such as sodium cyanoborohydride.[13–16] Hydrogen in the presence of a catalyst may also be employed as a reducing agent and is in fact the general method used to convert nitriles to amines, as in the second case.[13]

While there are various methods to introduce amines onto a scaffold, most would not be suitable for use with an indole system since reducing agents and high temperatures and pressures, as briefly discussed above, risk reducing, or even destroying, the indole scaffold.[13] For this reason, a stable oxime intermediate was chosen, instead of the classic imine intermediate, which could be reduced with zinc dust and glacial acetic acid, as opposed to the more commonly used $LiAlH_4$.[17,18] This method has the further benefits that the stable intermediate may be purified, using column chromatography, prior to reduction and solid hydroxylamine hydrochloride is much easier to work with than gaseous ammonia.

Oximes, like the name suggests, belong to the class of imines and have a general formula R_1R_2C=N-OH.[17] The first oxime was synthesised by Meyer and Janny in 1882, a time in which the existence of oximes was questioned. They named this compound "Acetoximsäure" (acetoximic acid) (*Figure 3.4*).[19] Soon after, "Acetoxim" (acetoxime) followed (*Figure 3.4*).[20]

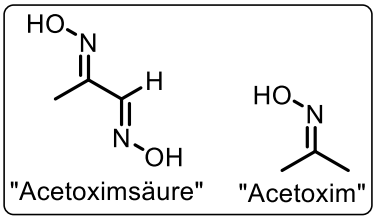
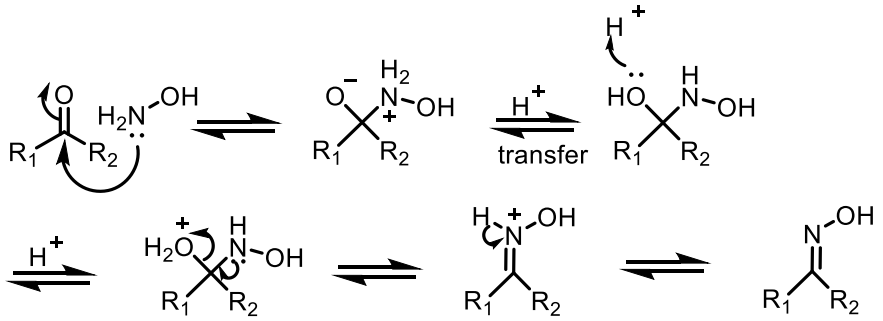


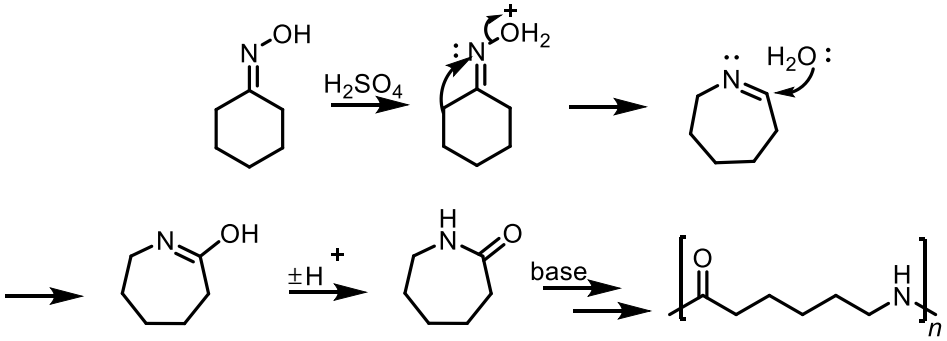
Figure 3.4: The first synthesised oximes.

Oxime synthesis generally involves the reaction of an aldehyde or ketone with hydroxylamine hydrochloride in the presence of a base. Hydroxylamine adds to the carbonyl group, through the more nucleophilic nitrogen atom, to form an unstable intermediate. The subsequent decomposition by loss of water results in an oxime (*Scheme 3.4*).[17]



Scheme 3.4: A general reaction mechanism for oxime formation.

An interesting industrial application of oximes is found in the manufacture of nylon. Nylon is industrially manufactured by the alkaline polymerization of a cyclic amide known as Caprolactam. Caprolactam is produced when the oxime of cyclohexanone undergoes a Beckmann rearrangement in the presence of sulfuric acid. In this reaction, the oxime is converted into a good leaving group which leaves upon alkyl migration onto the nitrogen. The cation is subsequently trapped by water to give the amide (*Scheme 3.5*). This amide, Caprolactam, is then used further for the synthesis of nylon.

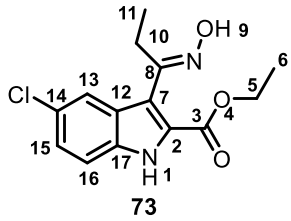


Scheme 3.5: The industrial manufacture of Nylon.

As in the case of this synthetic pathway, oximes, due to their stability, are often used as convenient intermediates in organic synthesis for the introduction of amines and amides.

3.2.3.1 Synthesis of ethyl 5-chloro-3-[1-(hydroxyimino)propyl]-1*H*-indole-2-carboxylate (73)[18] – Oxime formation

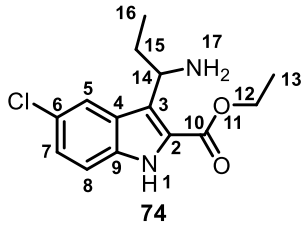
Oxime formation was achieved without difficulty by dissolving ethyl 5-chloro-3-propionyl-1*H*-indole-2-carboxylate (**72**) together with hydroxylamine hydrochloride and pyridine in THF. The highest yield was obtained when the reaction mixture was heated at reflux for 60 hours and a few activated molecular sieves were added to remove water produced as a byproduct, thereby driving the equilibrium towards the desired product. Furthermore, the solvent, THF, and pyridine were distilled and pre-dried over activated molecular sieves to keep the water in the reaction mixture to a minimum. The use of a Dean-Stark apparatus, filled with activated molecular sieves to draw water molecules over at the lower temperature of 85 °C, was found to be less effective than when the molecular sieves were added directly to the flask. After cooling the reaction mixture to room temperature, it was diluted with water, acidified with 1 M HCl and extracted with diethyl ether. Purification by column chromatography yielded the pure title product in 76% yield.



73

One of the most reliable methods to confirm the presence of an oxime is the combined use of NMR and IR spectroscopy. H_9 was clearly visible as a singlet integrating for 1 proton at 11.06 ppm in the ^{1}H NMR spectrum, while the ^{13}C NMR spectrum showed a clear shift of the C_8 signal from 201.4 ppm for the ketone to 154.8 ppm for the oxime, as expected from the greater shielding due to the electron density around the C=N in the oxime. Furthermore, from the IR spectrum, three oxime peaks could be identified: an O-H stretch at 3309 cm^{-1}, a C-N stretch at 1245 cm^{-1} and a N-O stretch at 953 cm^{-1}. The molecular ion was not confirmed with HRMS, but the correct structure was confirmed with NMR and IR spectroscopy and from the confirmed structures that followed.

3.2.3.2 Synthesis of ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (74)[18] – Reduction



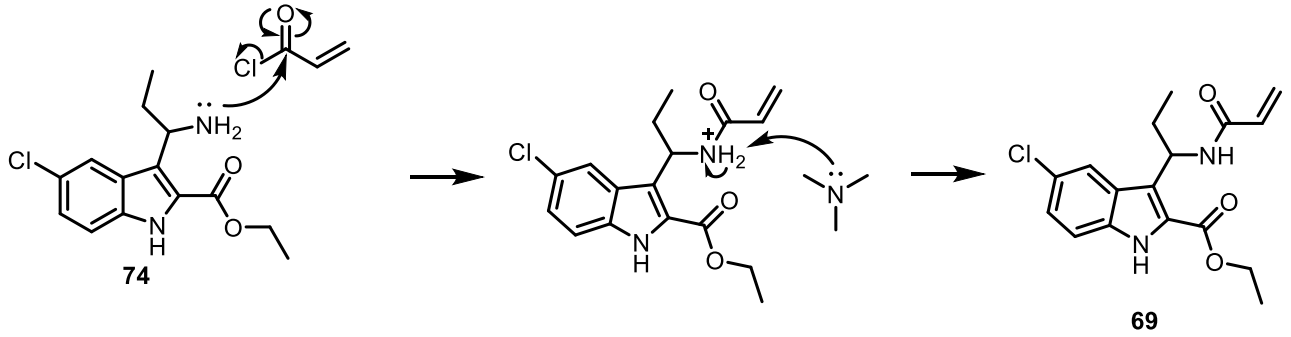
74

For the next reduction step, ethyl 5-chloro-3-[1-(hydroxyimino)propyl]-1*H*-indole-2-carboxylate (**73**) and sodium acetate were dissolved in glacial acetic acid after which zinc dust was added. The reaction mixture was stirred for 18 hours at room temperature and was then poured into a beaker containing ice cooled sodium hydroxide. The crude product was extracted with diethyl ether and was used in the next step without further purification. The ^{1}H NMR spectrum of the crude product confirmed the presence of amine protons H_{17}.

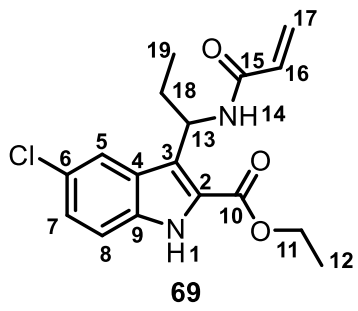
3.2.4 Introducing a moiety for irreversible binding in the binding pocket

3.2.4.1 Synthesis of ethyl 3-(1-acrylamidopropyl)-5-chloro-1*H*-indole-2-carboxylate (69)

The acrylamide moiety was next introduced by reacting crude ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (**74**) with acrolyl chloride in dichloromethane (DCM). Due to the exothermic nature of the reaction, a low temperature was maintained by means of an ice bath and acrolyl chloride was added dropwise. Trimethylamine was used to "mop up" the acid by-product as shown in *Scheme 3.6*. After general work-up procedures using ammonium chloride to quench the reaction and ethyl acetate for product extraction, purification by column chromatography resulted in an 85% yield of the desired product over the two steps from the oxime.



*Scheme 3.6: Reaction of acrolyl chloride and the indole amine **74** to form the target amide **69**.*



The product was easily identified by means of NMR spectroscopy. H_{14} was observed as a doublet at 8.59 ppm, H_{16} as a multiplet at 6.37 – 6.31 ppm and H_{17a} and H_{17b} as doublets of doublets at 5.99 ppm and 5.55 ppm, each of the signals integrating for 1 proton (*Figure 3.5*, circled in purple). Interestingly, by inspection of the *J* values, H_{17a} could be identified as *trans* to H_{16} and H_{17b} as *cis*. The larger *J* value of 17.02 Hz corresponds to the proton positioned *trans* to H_{16}, while the *J* value of 10.16 Hz corresponds to the *cis* proton. From the ^{13}C NMR spectrum, C_{15} was seen furthest downfield at 164.1 ppm. Furthermore, the $[M+H]^+$ ion was identified by high resolution mass spectrometry (HRMS) and corresponded well to the expected mass.

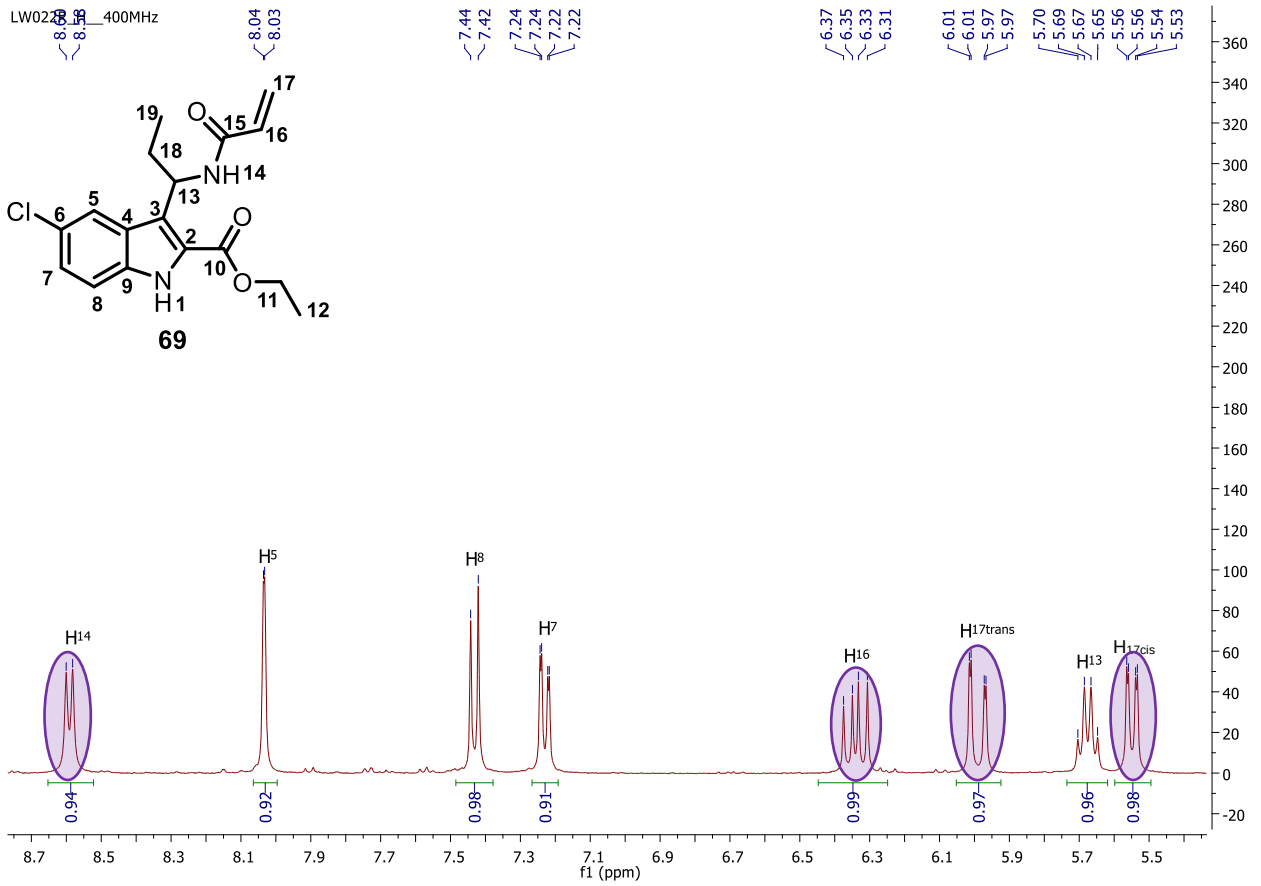
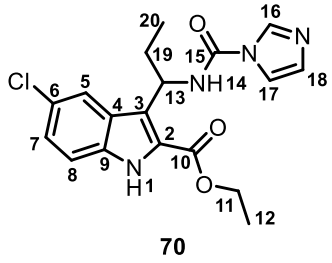


*Figure 3.5: The ^{1}H NMR spectrum of product **69** shown from 8.7 ppm to 5.4 ppm. From left to right the signals correspond to H_{14}, H_5, H_8, H_7, H_{16}, $H_{17trans}$, H_{13} and H_{17cis} as labelled. The proton signals corresponding to the acrylamide moiety (H_{14}, H_{16}, $H_{17trans}$ and H_{17cis}) are circled in purple.*

3.2.4.2 Synthesis of ethyl 3-[1-(1*H*-imidazole-1-carboxamido)propyl]-5-chloro-1*H*-indole-2-carboxylate (70)[21]

Similar to the reactions described in Section 2.3.7, 1,1'-carbonyldiimidazole (CDI) could react with the amine, ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (**74**), and imidazole could be eliminated. To this end, crude ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (**74**) was dissolved in THF and this solution was added dropwise to a second solution of CDI in THF. After 15 minutes, the reaction was complete and a work-up was done. Purification by column chromatography resulted in a 29% yield of the title compound over two steps. The low yield compared to **69** may be accounted for by the poor reproducibility of the oxime reduction. Nevertheless, the desired product was obtained and full characterisation could be done.



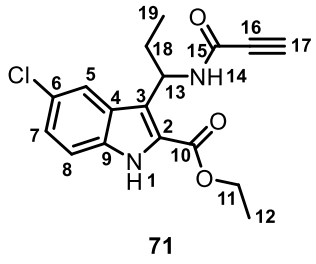
70

H_{14} was observed as a doublet at 8.82 ppm, integrating for 1 proton, indicating that the carbonyl imidazole moiety had been attached at this position. A multiplet at 8.29 – 8.22 ppm integrating for 1 proton was identified as H_{16}, while H_{17} and H_{18}, both multiplets integrating for 1 proton each, were located at 7.77 – 7.70 ppm and 7.06 – 6.98 ppm. The signal for C_{15} was located at 148.1 ppm, just slightly upfield from C_{10}. For final verification of the structure, the $[M-H]^-$ ion was confirmed by HRMS.

3.2.4.3 Synthesis of ethyl 5-chloro-3-(1-propiolamidopropyl)-1*H*-indole-2-carboxylate (71)

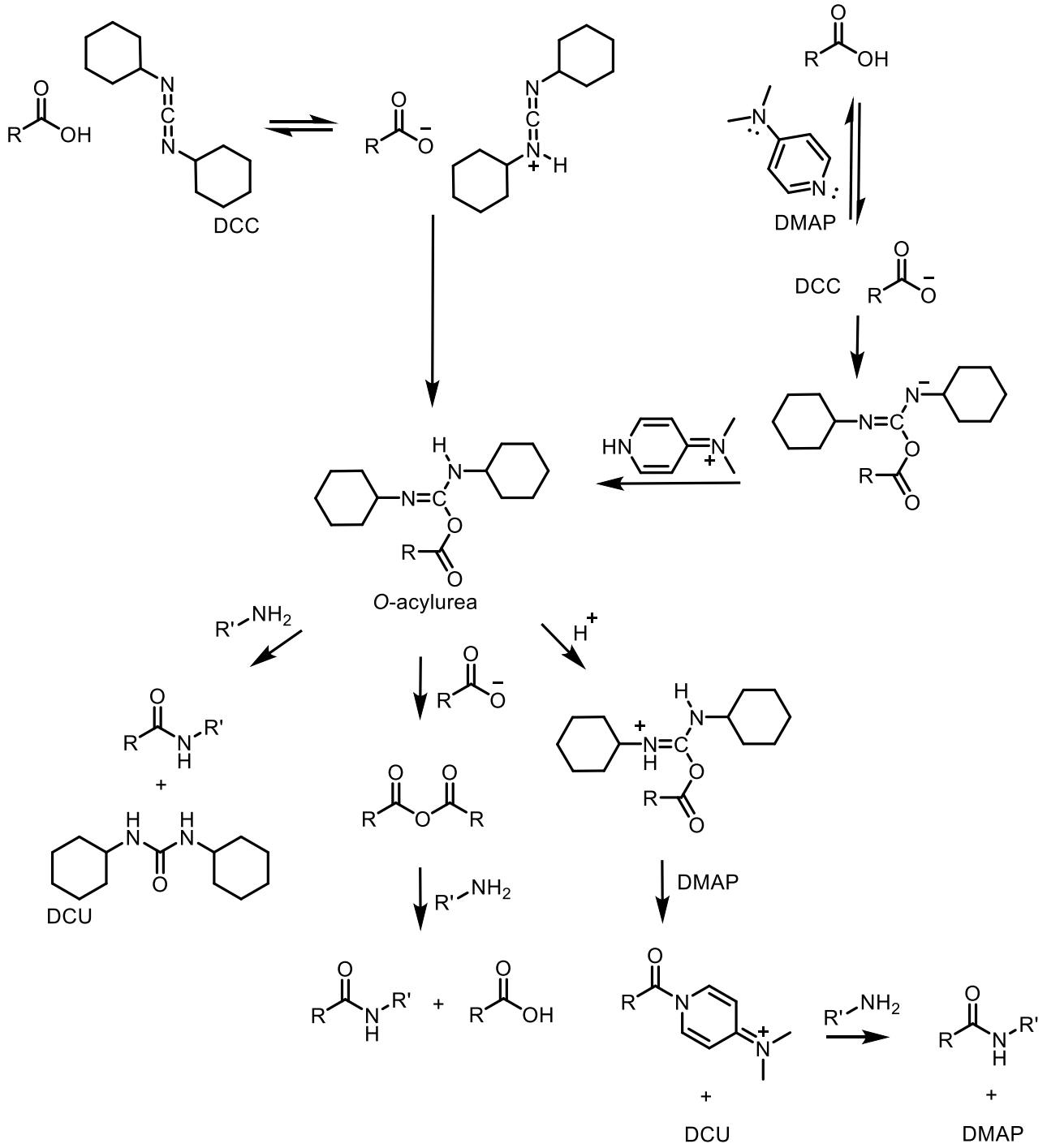
N,N'-Dicyclohexylcarbodiimide (DCC) is a coupling agent that has been widely used since 1955 to yield amides from carboxylic acids.[22,23] The first step is the reaction of the carboxylic acid with DCC to form an *O*-acylurea (*Scheme 3.7*). This intermediate may then form the amide *via* either direct coupling with the amine, producing dicyclohexylurea (DCU) as by-product, or by going through the carboxylic acid anhydride which will yield the amide when reacted with the amine.[22] A reaction adapted from the Steglich esterification reaction may also be used to obtain the desired amide.[24] In this version of the reaction, *O*-acylurea is reacted with 4-dimethylaminopyridine (DMAP) so as to install a better leaving group. The amine may then attack at the carbonyl carbon to replace DMAP.

The last mentioned synthetic method was used to synthesise compound **71**. The crude amine **74** was dissolved in DCM and the solution was subsequently cooled to 0 °C. Propiolic acid, DCC and DMAP were then added and the reaction mixture was allowed to stir at room temperature for three and a half hours. Upon completion, the solution was filtered and concentrated and the product was purified by column chromatography. A yield of 46% was obtained over the two steps from the oxime.



71

In terms of characterisation, from the ^{1}H NMR spectrum of the product, a doublet at 9.31 ppm, identified as H_{14}, and a singlet at 4.13 ppm, identified as H_{17}, confirmed the presence of the propiolamide moiety. Furthermore, the C_{15} signal was present at 151.6 ppm in the ^{13}C NMR spectrum. Further proof of the target compound was the $[M-H]^-$ ion identified by HRMS as 331.0849 ($C_{17}H_{16}ClN_2O_3$ requires 331.0859).



Scheme 3.7: From carboxylic acid to amide using DCC.

3.2.5 Biological testing results

With the three proof of concept compounds in hand, biological evaluation was carried out by our collaborator, Dr. Adriaan Basson, at the HIV Pathogenesis Research Unit at the University of the Witwatersrand, Medical School. To this end, an *in vitro* single-cycle, non-replicative phenotypic assay, using a HIV-1 retroviral vector system for the production of virus-like particles, was utilised.[25] The compounds were incubated with 293T cells together with the virus-like particles. After 48 hours, inhibition was quantified by luminescence measurement. The toxicity and activity assay results are

presented in *Table 3.1* below. DMSO showed no significant toxicity over the concentrations tested and nevirapine had an IC_{50} value within the expected range. The compounds were unfortunately found to be toxic and thus inactive as reversible inhibitors for HIV-1.

Table 3.1: A summary of the toxicity and activity data for the three proof of concept compounds.

Compounds	Toxicity (CC_{50}, μM)				Activity (IC_{50}, μM)			
	1	2	Ave	SD	1	2	Ave	SD
69	4.4	7.6	6.0	2.2	1.934	2.200	2.067	0.188
70	27.4	67.1	47.3	28.1	31.369	56.896	44.132	18.050
71	8.3	10.1	9.2	1.3	8.0651	10.679	9.655	1.434
NVP ctrl.					0.096	0.101	0.099	0.004
DMSO	>1%	>1%	>1%					

Ave = average

SD = standard deviation

NVP ctrl. = Nevirapine control

The toxicity and activity curves for each of the compounds were plotted to aid with the interpretation of the results (*Figure 3.6*). These showed that the concentration required for viral inhibition was not viable for cell life. The cells were effectively killed at the same concentration where the virus was inhibited. Ideally, the concentration at which the virus is inhibited should be much lower than the concentration at which cell viability decreases. One possible reason for the toxicity observed is off-target inhibition resulting in cell death.

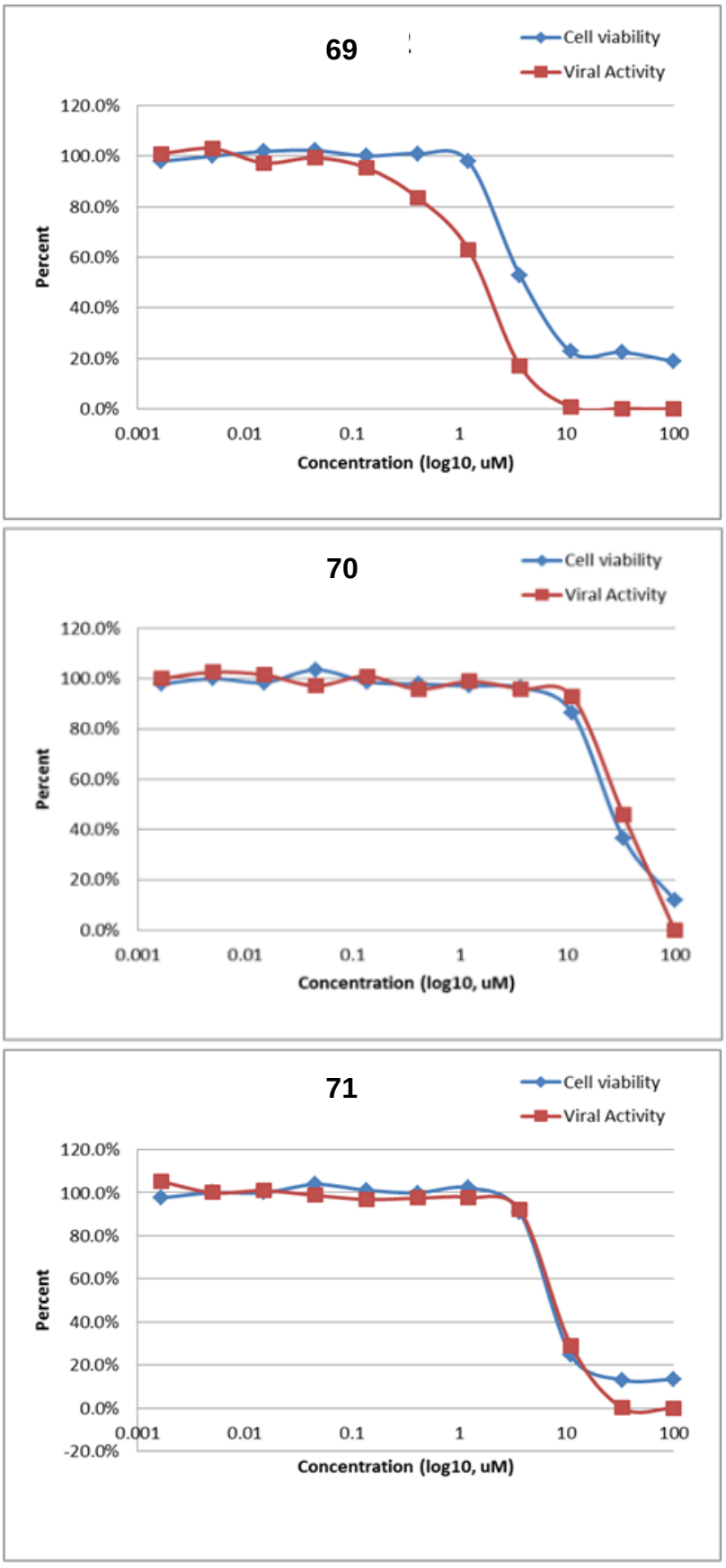
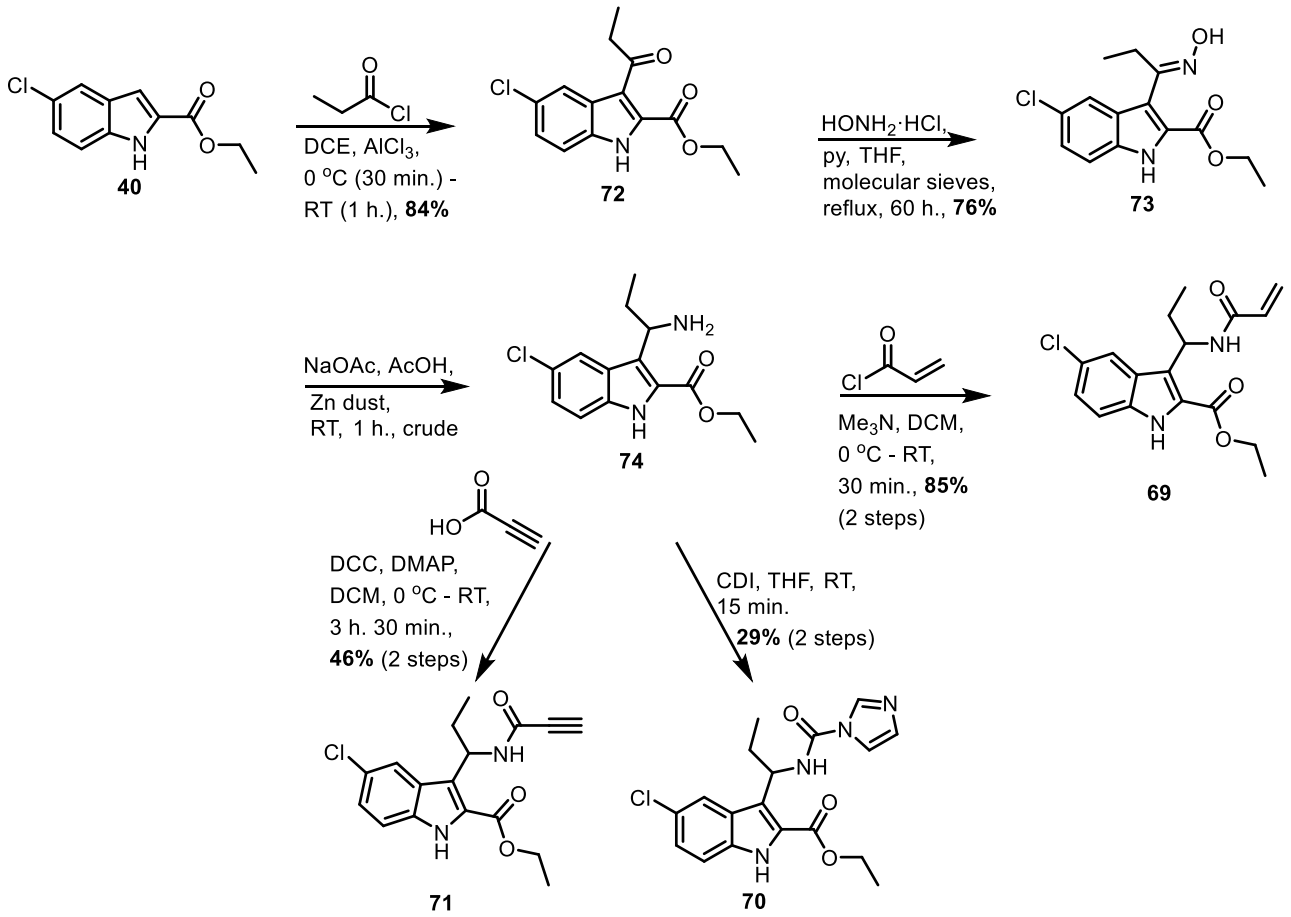


*Figure 3.6: The toxicity and activity curves for compounds **69**, **70** and **71**.*

3.2.6 Concluding remarks

In our endeavour to synthesise the first covalent inhibitors of the Y181C mutant strain of the HI virus, a successful reversible inhibitor scaffold, designed extensively using molecular modelling, was selected for modification and a literature search for appropriate electrophilic warheads was conducted. The indole base-scaffold was to be fitted with an acrylamide, an imidazole carboxamide and a propiolamide moiety so as to synthesise three proof of concept compounds for biological evaluation. The successful synthesis of the target compounds is summarised in *Scheme 3.8*. After full characterisation, these compounds were sent for biological evaluation. Unfortunately, compounds **69**, **70** and **71** proved to be cytotoxic. Careful optimization of the non-covalent binding affinity, by molecular modelling and trial and error synthesis, would have to be undertaken to develop irreversible inhibitors that could be of use. Such a lengthy process was, however, beyond the scope of this MSc project and will be considered for future work (see Section 6.2).



Scheme 3.8: A summary of the synthesis of three proof of concept compounds.

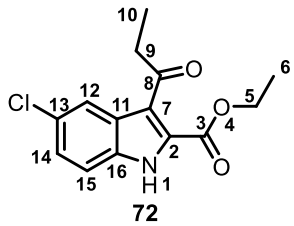
3.3 Experimental section

Please refer to Section 2.4.1, Chapter 2 for general procedures pertaining to the purification of reagents and solvents, chromatography, spectroscopic and physical data, molecular modelling and

other general procedures followed. Note that an arbitrary numbering system was used and reference to the provided structure should be made when studying the NMR spectra.

3.3.1 Synthesis of ethyl 5-chloro-3-propionyl-1*H*-indole-2-carboxylate (72)

Propionyl chloride (1.16 mL, 13.3 mmol) was dissolved in 1,2-dichloroethane (DCE) (20 mL) in a two-neck round-bottom flask, which had been flushed with alternating vacuum and nitrogen purge cycles, after which anhydrous aluminium chloride ($AlCl_3$) (1.77 g, 13.3 mmol) was added at 0 °C. The reaction mixture was left to stir on ice for 30 min. Ethyl 5-chloro-1*H*-indole-2-carboxylate (**40**) (1.00 g, 4.47 mmol) was then added and the reaction mixture was stirred at RT for 1 h. Upon completion, ethyl acetate (EtOAc) (30 mL) was added to the reaction mixture which was then decanted into a separating funnel containing $NaHCO_3$ (100 mL). Extraction was done with EtOAc (3 x 30 mL) and the combined organic layers were washed with brine (30 mL), dried over magnesium sulfate ($MgSO_4$), filtered and evaporated under reduced pressure. Purification was done by column chromatography (5 - 50% EtOAc/Hexane) to yield the title compound (1.04 g, 3.74 mmol, 84%) (R_f = 0.27, 20% EtOAc/Hexane) as a white solid.



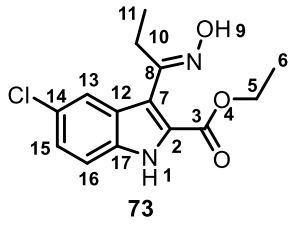
72

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.12 (s, 1H, H_1), 7.97 (d, *J* = 1.9 Hz, 1H, H_{12}), 7.35 (dd, *J* = 8.7, 0.5 Hz, 1H H_{15}), 7.31 (dd, *J* = 8.7, 1.9 Hz, 1H, H_{14}), 4.46 (q, *J* = 7.2 Hz, 2H, H_5), 3.07 (q, *J* = 7.3 Hz, 2H, H_9), 1.43 (t, *J* = 7.2 Hz, 3H, H_6), 1.23 (t, *J* = 7.3 Hz, 3H, H_{10}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 201.4 (C_8), 160.6 (C_3), 133.3, 128.6, 127.8, 127.0, 126.9, 122.2, 121.5, 112.9, 62.3 (CH_2), 37.4 (CH_2), 14.4 (CH_3), 8.7 (CH_3).

Nuclear magnetic resonance characterization compared well to data in the literature.[12]

3.3.2 Synthesis of ethyl 5-chloro-3-[1-(hydroxyimino)propyl]-1*H*-indole-2-carboxylate (73)

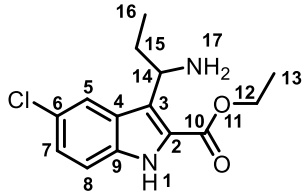
To a 20 mL two-neck round-bottom flask fitted with a condenser, which had been flushed with alternating vacuum and nitrogen purge cycles, ethyl 5-chloro-3-propionyl-1*H*-indole-2-carboxylate (**72**) (0.200 g, 0.715 mmol), hydroxylamine hydrochloride (0.120 g, 1.75 mmol), dry tetrahydrofuran (THF) (10 mL), dry pyridine (2.60 mL, 23.1 mmol) and approximately 10 activated molecular sieves were added. The reaction mixture was then heated at reflux for 60 h. Upon completion the cooled reaction mixture was diluted with water (40 mL), acidified with HCl (60 mL, 1 M) and extracted with diethyl ether (3 x 30 mL). The combined organic layers were washed with water, dried over $MgSO_4$, filtered and evaporated under reduced pressure. Purification was done by column chromatography (5 - 50% EtOAc/Hexane) to yield the title compound (0.160 g, 0.543 mmol, 76%) (R_f = 0.14, 20% EtOAc/Hexane) as a white solid.



73

MP 166 – 170 °C. **IR (ATR, cm^{-1})** 3309 (O-H str), 2985 (C-H str), 2941 (C-H str), 1689 (C=O str), 1546, 1462, 1245 (C-N str), 1181, 953 (N-O str), 901, 861, 801, 780.**^{1}H NMR (300 MHz, DMSO-d) δ** 12.16 (s, 1H, H_1), 11.06 (s, 1H, H_9), 7.55 (d, *J* = 2.0 Hz, 1H H_{13}), 7.48 (d, *J* = 8.8 Hz, 1H, H_{16}), 7.29 (dd, *J* = 8.8, 2.0 Hz, 1H, H_{15}), 4.32 (q, *J* = 7.1 Hz, 2H, H_5), 2.72 (q, *J* = 7.5 Hz, 2H, H_{11}), 1.31 (t, *J* = 7.1 Hz, 3H, H_6), 0.88 (t, *J* = 7.5 Hz, 3H, H_{10}). **^{13}C NMR (101 MHz, DMSO-d) δ** 160.7 (C_3), 154.8 (C_8), 134.4, 128.2, 125.8, 125.2, 120.0, 117.5, 114.4, 60.9, 30.7 (CH_2), 22.4 (CH_2), 14.1 (CH_3), 10.4 (CH_3).

3.3.3 Synthesis of ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (74)

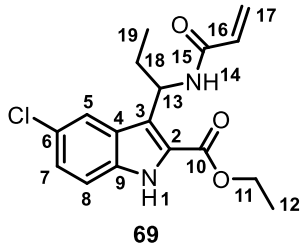


74

Ethyl 5-chloro-3-[1-(hydroxyimino)propyl]-1*H*-indole-2-carboxylate (**73**) (0.105 g, 0.356 mmol) and sodium acetate (0.605 g, 7.37 mmol) were dissolved in glacial acetic acid (10 mL) in a 20 mL two-neck round-bottom flask. After the addition of zinc dust (0.356 g, 5.44 mmol), the mixture was stirred at RT for 18 h, poured into an ice-cooled NaOH solution (200 mL, 2 M) and extracted with diethyl ether (3 x 50 mL). The organic layers were dried over $MgSO_4$, filtered and evaporated under reduced pressure. The crude product, obtained as a white solid, was used without further purification in the next reaction.

3.3.4 Synthesis of ethyl 3-(1-acrylamidopropyl)-5-chloro-1*H*-indole-2-carboxylate (69)

Crude ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (**74**) (0.100 g, 0.356 mmol) and trimethylamine (Me_3N) (0.06 mL, 0.4 mmol) were dissolved in dichloromethane (DCM) (5 mL) in a 20 mL two-neck round-bottom flask, fitted with a rubber septum, which had been flushed with alternating vacuum and nitrogen purge cycles, and the mixture was cooled to 0 °C. Acrolyl chloride (0.03 mL, 0.4 mmol) in DCM (1 mL) was then added dropwise over 5 min. The reaction mixture was allowed to stir for 30 min, warming up to RT. NH_4Cl (50 mL) was then added and extraction was done with EtOAc (3 x 30 mL). The combined organic layers were dried over $MgSO_4$ and the solvent was removed under reduced pressure. The title compound [0.102 g, 0.304 mmol, 85% (over two steps)] (R_f = 0.41, 50% EtOAc/Hexane) was obtained as a white solid after purification with column chromatography (5 - 50% EtOAc/Hexane).

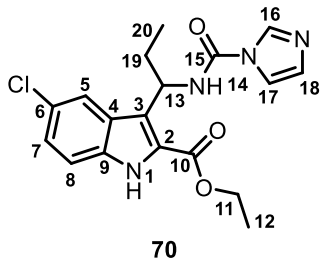


69

MP 144 – 146 °C. **IR (ATR, cm^{-1})** 3242 (N-H str), 2966 (C-H str), 1694 (C=O str), 1657 (C=O str), 1537, 1440, 1247 (C-O str), 1098, 802, 781. **^{1}H NMR (400 MHz, DMSO-d) δ** 11.79 (s, 1H, H_1), 8.59 (d, *J* = 7.2 Hz, 1H, H_{14}), 8.04 (d, *J* = 1.6 Hz, 1H, H_5), 7.43 (d, *J* = 8.8 Hz, 1H, H_8), 7.23 (dd, *J* = 8.8, 1.6 Hz, 1H, H_7), 6.37 – 6.31 (m, 1H, H_{16}), 5.99 (dd, *J* = 17.0, 2.0 Hz, 1H, $H_{17trans}$), 5.68 (q, *J* = 7.6, 1H, H_{13}), 5.55 (dd, *J* = 10.2, 2.0 Hz, 1H, H_{17cis}), 4.44 – 4.28 (m, 2H, CH_2), 2.00 – 1.71 (m, 2H, CH_2), 1.36 (t, *J* = 7.0 Hz, 3H, CH_3), 0,87 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, DMSO-d) δ** 164.1 (C_{15}), 161.3 (C_{10}), 134.8, 131.7, 125.9, 125.3, 124.7, 124.5, 124.1, 123.8, 120.6, 114.4, 60.7, 47.6, 28.1, 14.3 (CH_3), 11.2 (CH_3). **HRMS** calculated for $C_{17}H_{20}ClN_2O_3$ $[M+H]^+$, 335.1166, found 335.1166.

3.3.5 Synthesis of ethyl 3-[1-(1*H*-imidazole-1-carboxamido)propyl]-5-chloro-1*H*-indole-2-carboxylate (70)

A solution of crude ethyl 3-(1-aminopropyl)-5-chloro-1*H*-indole-2-carboxylate (**74**) (0.122 g, 0.436 mmol) in anhydrous THF (5 mL), under nitrogen atmosphere, was added dropwise to a stirred suspension of 1,1'-carbonyldiimidazole (CDI) (0.0848 g, 0.523 mmol) in anhydrous THF (5 mL) in a 20 mL two-neck round-bottom flask, which had been flushed with alternating vacuum and nitrogen purge cycles, at RT. After addition, the mixture was stirred for a further 15 min. The solvent was then evaporated. The residue was re-dissolved in DCM (10 mL) and washed with water (2 x 5 mL). The organic layer was dried over $MgSO_4$ and filtered after which the solvent was evaporated under reduced pressure. Purification with column chromatography (20 - 80% EtOAc/Hexane) yielded the title compound [0.0740 g, 0.198 mmol, 29% (over two steps)] (R_f = 0.22, 70% EtOAc/Hexane) as a white solid.

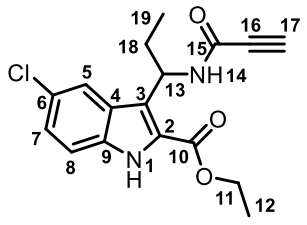


70

MP 114 – 116 °C. **IR (ATR, cm^{-1})** 3586, 3255 (N-H str), 1710 (C=O str), 1694, 1547 (N-H bend), 1478, 1438, 1373, 1323, 1290, 1242 (C-N str), 1193, 1098, 1074, 1014. **^{1}H NMR (300 MHz, DMSO-d) δ** 11.88 (s, 1H, H_1), 8.82 (d, *J* = 6.8 Hz, 1H, H_{14}), 8.29 – 8.22 (m, 1H, H_{16}), 8.15 (d, *J* = 2.0 Hz, 1H, H_5), 7.77 – 7.70 (m, 1H, $H_{17/18}$), 7.45 (d, *J* = 8.8 Hz, 1H, H_8), 7.24 (dd, *J* = 8.8, 2.0 Hz, 1H, H_7), 7.06 – 6.98 (m, 1H, $H_{17/18}$), 5.77 – 5.69 (m, 1H, H_{13}), 4.39 (q, *J* = 7.1 Hz, 2H, H_{11}), 2.27 – 1.77 (m, 2H, H_{19}), 1.38 (t, *J* = 7.1 Hz, 3H, H_{12}), 0.94 (t, *J* = 7.3 Hz, 3H, H_{20}). **^{13}C NMR (75 MHz, DMSO-d) δ** 161.3 (C_{10}), 148.1 (C_{15}), 136.0, 134.7, 129.5, 125.6, 124.9, 124.8, 124.4, 122.6, 120.4, 116.7, 114.5, 60.7, 49.7, 27.4, 14.2 (CH_3), 11.3 (CH_3). **HRMS** calculated for $C_{18}H_{18}ClN_4O_3$ $[M-H]^-$, 373.1066, found 373,1067.

3.3.6 Synthesis of ethyl 5-chloro-3-(1-propiolamidopropyl)-1*H*-indole-2-carboxylate (71)

Crude ethyl 3-(1-aminopropyl)-5-chloro-1H-indole-2-carboxylate (**74**) (0.125 g, 0.444 mmol) was dissolved in dry DCM (1 mL) and cooled to 0 °C in a 5 mL two-neck round-bottom flask, which had been flushed with alternating vacuum and nitrogen purge cycles. Propiolic acid (0.03 mL, 0.4 mmol), *N,N'*-dicyclohexylcarbodiimide (DCC) (0.110 g, 0.533 mmol) and 4-dimethylaminopyridine (DMAP) (0.00540 g, 0.0444 mmol) were then added. The mixture was stirred at RT for 3 h. and 30 min. after which the solution was filtered and concentrated. The crude product was purified by column chromatography (5 - 50% EtOAc/Hexane) to yield the title compound [0.105 g, 0.314 mmol, 46% (over two steps)] (R_f = 0.48, 50% EtOAc/Hexane) as a white solid.



71

MP 138 – 140 °C. **IR (ATR, cm^{-1})** 3283 (N-H str), 2162 (C≡C str), 1691 (C=O str), 1649 (C=O str), 1525, 1437, 1252 (C-N str), 1193, 890, 808. **^{1}H NMR (400 MHz, DMSO-d) δ** 11.81 (s, 1H, H_1), 9.31 (d, *J* = 7.5 Hz, 1H, H_{14}), 8.11 (d, *J* = 1.9 Hz, 1H, H_5), 7.44 (d, *J* = 8.8 Hz, 1H, H_8), 7.25 (dd, *J* = 8.8, 1.9 Hz, 1H, H_7), 5.65 (dd, *J* = 15.4, 7.5 Hz, 1H, H_{13}), 4.36 (q, *J* = 7.0 Hz, 2H, H_{11}), 4.13 (s, 1H, H_{17}), 1.96 – 1.73 (m, 2H, H_{18}), 1.35 (t, *J* = 7.0 Hz, 3H, H_{12}), 0.86 (t, *J* = 7.3 Hz, 3H, H_{19}). **^{13}C NMR (75 MHz, DMSO-d) δ** 161.7 (C_{10}), 151.6 (C_{15}), 135.2, 126.1, 125.2, 125.0, 124.7, 123.3, 121.1, 114.8, 78.8, 76.1, 61.1, 48.4, 28.0, 14.7 (CH_3), 11.6 (CH_3). **HRMS** calculated for $C_{17}H_{16}ClN_2O_3$ $[M-H]^-$, 331.0859, found 331,0849.

3.4 Bibliography

1 C. Gonzalez-Bello, *ChemMedChem*, 2016, **11**, 22–30.

2 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discov.*, 2011, **10**, 307–317.

3 A. H. Chan, W.-G. Lee, K. A. Spasov, J. A. Cisneros, S. N. Kudalkar, Z. O. Petrova, A. B. Buckingham, K. S. Anderson and W. L. Jorgensen, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 9725–9730.

4 M.-P. de Béthune, *Antiviral Res.*, 2010, **85**, 75–90.

5 P. A. J. Janssen, P. J. Lewi, E. Arnold, F. Daeyaert, M. De Jonge, J. Heeres, L. Koymans, M. Vinkers, J. Guillemont, E. Pasquier, M. Kukla, D. Ludovici, K. Andries, M. P. De Béthune, R. Pauwels, K. Das, A. D. Clark, Y. V. Frenkel, S. H. Hughes, B. Medaer, F. De Knaep, H. Bohets, F. De Clerck, A. Lampo, P. Williams and P. Stoffels, *J. Med. Chem.*, 2005, **48**, 1901–1909.

6 L. Petersen, C. H. Jessen, E. B. Pedersen and C. Nielsen, *Org. Biomol. Chem.*, 2003, **1**, 3541–3545.

7 Y. L. Aly, E. B. Pedersen, P. La Colla and R. Loddo, *Monatsh. Chemie*, 2006, **137**, 1557–1570.

8 M. Seki, Y. Sadakata, S. Yuasa and M. Baba, *Antivir. Chem. Chemother.*, 1995, **6**, 73–79.

9 G. M. Szczech, P. Furman, G. R. Painter, D. W. Barry, K. Borroto-Esoda, T. B. Grizzle, M. R. Blum, J.-P. Sommadossi, R. Endoh, T. Niwa, M. Yamamoto and C. Moxham, *Antimicrob. Agents Chemother.*, 2000, **44**, 123–130.

10 L. Petersen, E. B. Pedersen and C. Nielsen, *Synthesis*, 2001, **4**, 559–564.

11 W.-G. Lee, A. H. Chan, K. A. Spasov, K. S. Anderson and W. L. Jorgensen, *ACS Med. Chem. Lett.*, 2016, **7**, 1156–1160.

12 L. Khurana, H. I. Ali, T. Olszewska, K. H. Ahn, A. Damaraju, D. A. Kendall and D. Lu, *J. Med. Chem.*, 2014, **57**, 3040–3052.

13 S. Gomez, J. A. Peters and T. Maschmeyer, *Adv. Synth. Catal.*, 2002, **344**, 1037–1057.

14 R. Leuckart, *Ann. Chem. Pharm.*, 1885, **154**, 2341–2344.

15 K. A. Schellenberg, *J. Org. Chem.*, 1963, **28**, 3259–3261.

16 S. Kim, C. H. Oh, J. S. Ko, K. H. Ahn and Y. J. Kim, *J. Org. Chem.*, 1985, **50**, 1927–1932.

17 J. Clayden, N. Greevs, S. Warren and P. Wothers, *Organic Chemistry*, Oxford University Press, Oxford, Second., 2001.

18 M. Lehr, *Arch. Pharm.*, 1996, **329**, 386–392.

19 V. Meyer and A. Janny, *Berichte der Dtsch. Chem. Gesellschaft*, 1882, **15**, 1164–1167.

20 V. Meyer and A. Janny, *Berichte der Dtsch. Chem. Gesellschaft*, 1882, **15**, 1324–1326.

21 G. Gros, L. Martinez, A. S. Gimenez, P. Adler, P. Maurin, R. Wolkowicz, P. Falson and J. Hasserodt, *Bioorg. Med. Chem.*, 2013, **21**, 5407–5413.

22 E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606–631.

23 J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067–1068.

24 B. Neises and W. Steglich, *Angew. Chemie Int. Ed.*, 1978, **17**, 522–524.

25 M. Hassam, A. E. Basson, D. C. Liotta, L. Morris, W. A. L. van Otterlo and S. C. Pelly, *ACS Med. Chem. Lett.*, 2012, **3**, 470–475.

Chapter 4: Diaryl Ethers as Possible Non-nucleoside Reverse Transcriptase Inhibitors

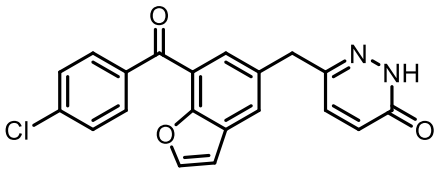
4.1 Our strategy

Herein the rationale behind the design and synthesis of diaryl ether non-nucleoside reverse transcriptase inhibitors are presented and a specific target compound is introduced.

4.1.1 Treasure hunt

In search for novel NNRTIs, a high-throughput screen of the Roche compound library yielded promising results with the discovery of pyridazinone **75** (*Table 4.1*), active against wild-type HIV-1 RT. In addition, this compound proved to be potent against two clinically relevant mutant viral strains namely K103N and Y181C.[1]

Table 4.1: A pyridazinone lead compound active against wild-type and mutant HIV-1 RT.[1]

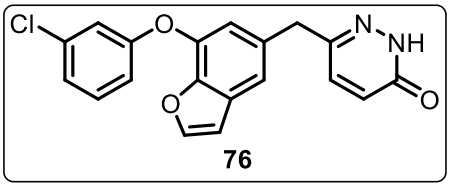


75

Assay	WT	K103N	Y181C
IC_{50} (μM)	0.40	0.26	0.36
EC_{50} (μM)	0.08	0.07	0.21

Using **75** as a lead molecule for their discovery program, Sweeney *et al.* substituted the ketone linker, between the terminal phenyl ring and the benzofuran ring, for an ether moiety, as an initial modification, and subsequently synthesised the potent diaryl ether analogue **76** (IC_{50} 1.0 μM versus the wild-type enzyme) (*Figure 4.1*). This lead compound was, however, found to be a uniformly strong inhibitor of CYP3A4 (IC_{50} 0.41 μM), an important enzyme in the digestive system.[1] Compounds that did not contain the benzofuran moiety, however, remained potent against wild-type HIV-1 and were relatively weak inhibitors of CYP3A4 (**77** and **78**, *Table 4.2*).[1] It was then found that fluorination of the central ring improved potency against both wild-type and mutant viral strains (*Table 4.3*). Crystallographic studies of a complex of **81** with the wild-type RT enzyme, suggested a possible binding mode of this series of inhibitors and revealed important interactions within the binding pocket. The pyridazinone ring formed a bidentate hydrogen-bonding interaction with the amide backbone of K103, the chlorine atom occupied a small pocket characterised by the side chains of V106 and V179, and the terminal phenyl ring engaged in hydrophobic interactions with Y188, Y181, and W229.[1]

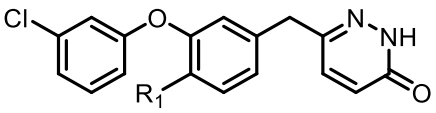
Further optimization through substitution of the terminal phenyl ring, lead to analogues with low nanomolar activity against wild-type and mutant viral strains (*Table 4.4*).[1]



76

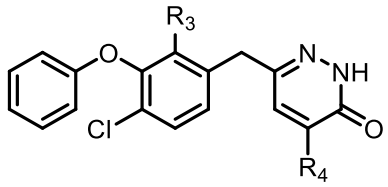
*Figure 4.1: Modifications to **75,** by Sweeney et al., lead to a potent diaryl ether analogue **76**.[1]*

Table 4.2: SAR of central phenyl ring substitution.[1]



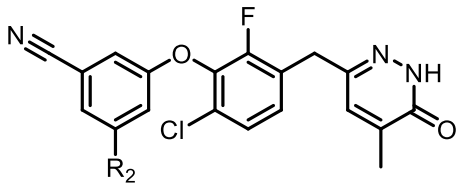
Compound	R_1	IC_{50} (μM)
77	Me	0.64
78	Cl	0.10

Table 4.3: Fluorination of the central ring improved potency [WT, K103N and Y181C refer to the IC_{50} values (μM) versus the enzyme indicated].[1]



Compound	R_3	R_4	WT	K103N	Y181C
79	H	H	0.52	3.35	5.49
80	H	Me	0.36	0.92	2.13
81	F	H	0.10	0.43	0.5
82	F	Me	0.02	0.05	0.21

Table 4.4: SAR of terminal phenyl ring substitution.[1]



Compound	R_2	WT	WT + serum	K103N	Y181C	G190A	K103N/L100I	K103N/Y181C
83	F	0.001	0.011	0.002	0.005	0.001	0.003	0.041
84	OMe	0.002	0.015	nt	0.009	0.001	0.002	0.031
85	CN	0.001	0.007	0.002	0.005	0.001	0.007	0.019

In order to improve the oral bioavailability of these compounds, the pyridazinone structure was modified. These efforts resulted in triazolinone, imidazolinone, imidazopyridinone, imidazopyridazinone, pyrazolopyridazine, hydantoin and uracil bioisosteres, with improved aqueous solubility and excellent activity against both wild-type and mutant strains of HIV-1.[2–4]

4.1.2 The ether moiety, now more common in NNRTIs

Shortly after the use of an ether moiety in NNRTI structures by Sweeney *et al.*, lersivirine was developed at the Pfizer Global Research and Development Laboratories and quickly progressed through clinical trials.[5,6] This compound resulted from the optimization of capravirine, a biaryl sulphide able to retain excellent activity against mutant strains of HIV (*Figure 4.2*).[7,8] Two or more RT mutations were required to achieve high level resistance to capravirine, but unfortunately, oxidative metabolism of this drug candidate into metabolites with much lower antiviral activity, resulted in rapid *in vivo* clearance and discouraged further clinical trials.[9] Lersivirine was designed to maintain the potency of its precursor and additionally display metabolic stability. A randomized, double-blind, phase IIb clinical trial in HIV-1-infected treatment-naïve patients, showed the antiviral efficacy of lersivirine to be comparable to that of efavirenz, but with fewer side-effects over a period of 24 to 48 weeks.[5] Furthermore, lersivirine was found to inhibit over 60% of mutant viruses and was selective against a range of human targets.[6] Similar to compound **81** (*Table 4.3*), lersivirine formed interactions with residues V106, Y181, Y188 and W229, among others interactions within the RT binding pocket.[6]

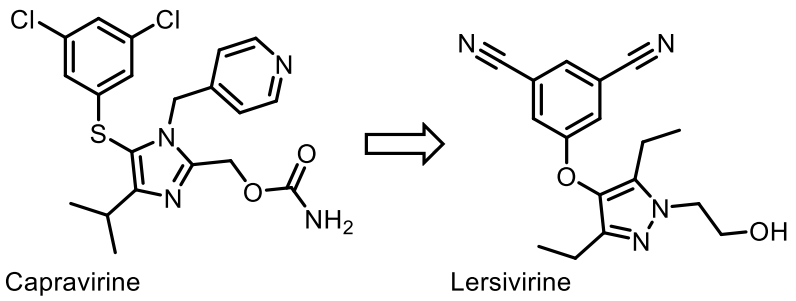


Figure 4.2: The lead optimization of capravirine resulted in lersivirine.[5,6,8]

Further examples of diaryl ether NNRTIs, with exceptional antiviral activity against both wild-type and mutant strains, including the highly prevalent K103N and Y181C mutations, include MK-1107, MK-4965 and MK-1439 (*Figure 4.3*). The interactions with the Y181, Y188 and W229 residues were conserved, as with the other diaryl ether NNRTIs discussed.[10,11]

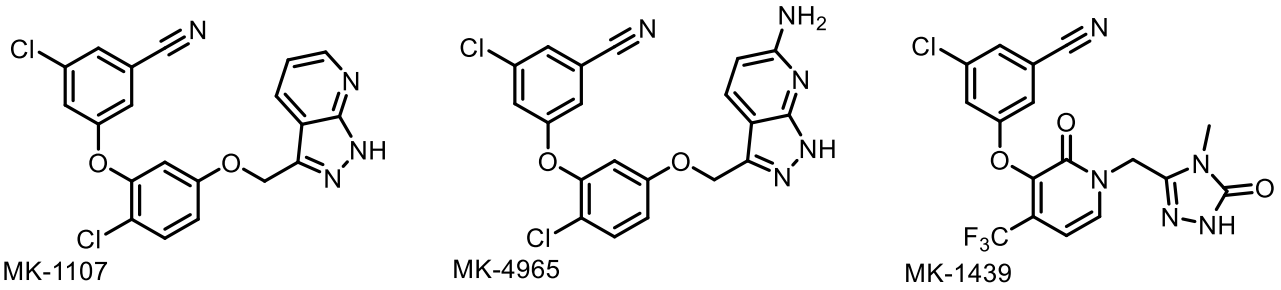


Figure 4.3: Further examples of the diaryl ether scaffold in potent NNRTIs.[10,11]

4.1.3 Designing a superior NNRTI

The detailed analysis of multiple crystal structures of NNRTIs bound in the RT enzyme, as well as data on drug resistant mutations, have resulted in the identification of key features necessary for improved resistance profiles. Among the important traits are conformational flexibility, positional adaptability and the targeting of highly conserved residues in the NNRTI-BP.[12] Drawing inspiration from the successes discussed previously, we sought to design and synthesise a superior NNRTI, with both the required flexibility and desirable interactions with residues in the NNRTI-BP.[1–4,6–8,10,11,13,14] It was envisaged that the flexible ether moiety would allow the inhibitor to compensate for the effects of mutation by avoiding mutated amino acids. From the four residues lining the NNRTI-BP, that have been identified as highly conserved namely: F227, W229, L234 and Y318, W229 was identified as a prime candidate for targeted design, mainly because it is a key component of the characteristic NNRTI-BP.[12,15] Interactions with this residue are common among the known diaryl ether NNRTIs and it was thus selected as valid target for our design.[1,6,10,11] The possibility of targeting Tyr318, a second conserved amino acid, through π-π interactions with a second aromatic ring, was noted and a diaryl ether base-scaffold unfolded (*Figure 4.4*). Molecular modelling was then used to elaborate the base-scaffold. The often present and successful chloro-benzonitrile moiety was selected as half of the structure (*Figure 4.4*, left-hand side).[10,11,13] It was foreseen that a methoxy group on the second aromatic ring, could occupy the valine 179 pocket and that a correctly positioned

amine could form hydrogen bonding interactions with lysine 101. Furthermore, a heteroatom, chlorine, was incorporated in the *ortho* position, based on previous successes (*Figure 4.4*, right-hand side).[1–4]

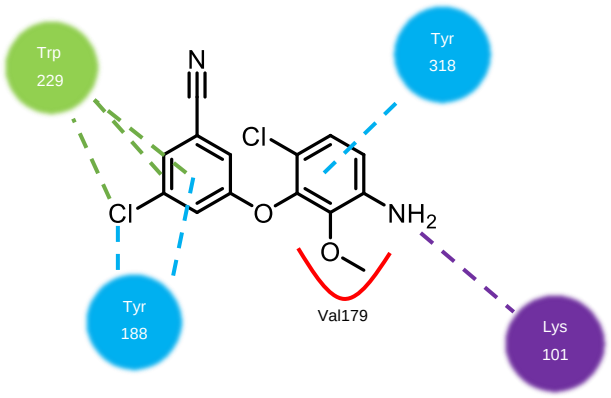


Figure 4.4: The interactions envisioned for a designed diaryl NNRTI in the NNRTI-BP.

4.2 A two-part synthesis

4.2.1 A proposed synthetic route

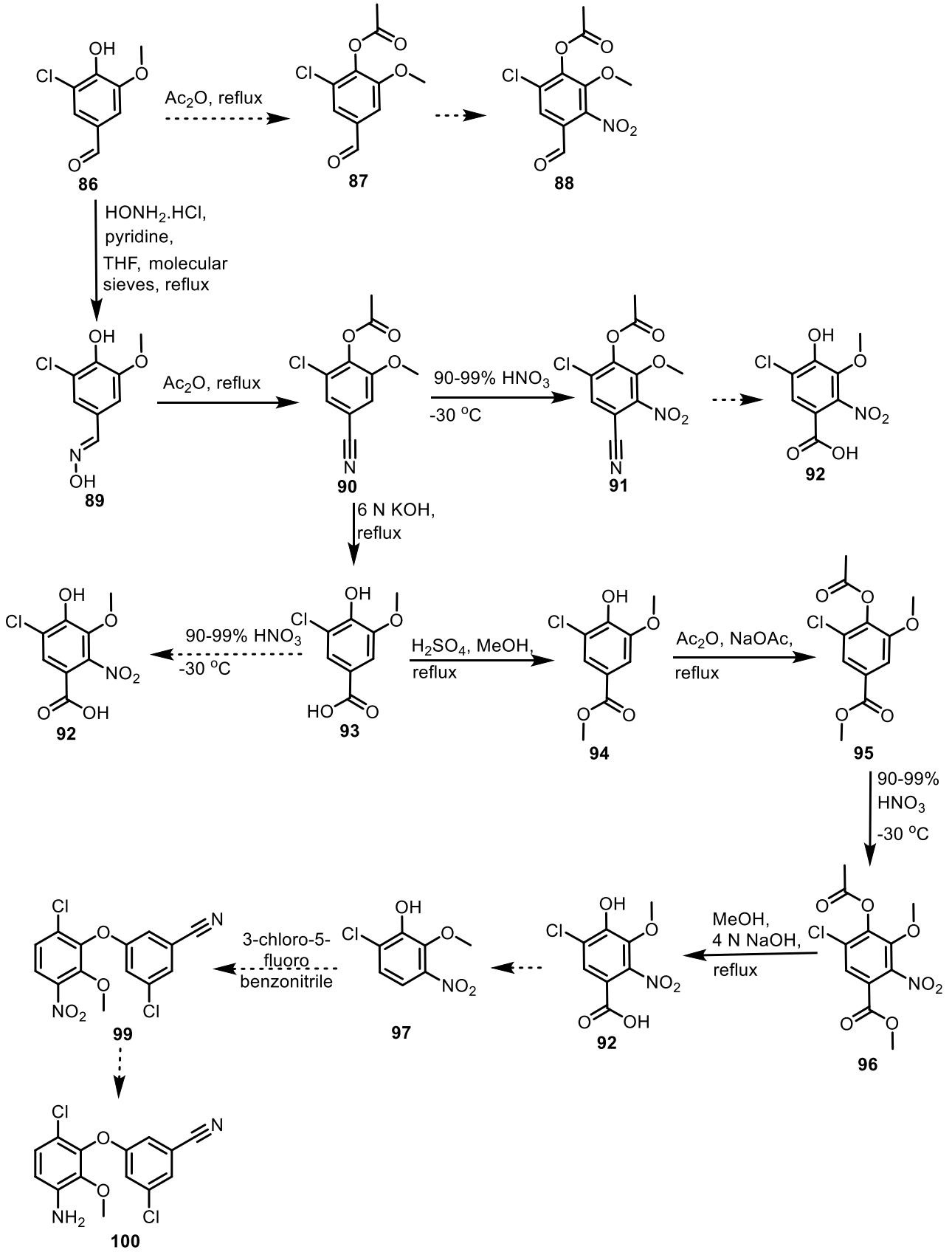
It was proposed that the two substituted aromatic rings could be obtained separately, after which nucleophilic aromatic substitution (S_NAr) could be employed to link the rings, through an ether moiety, to obtain the target compound 3-(3-amino-6-chloro-2-methoxyphenoxy)-5-chlorobenzonitrile (**89**, *Scheme 4.1*). In terms of a brief summary, Raiford & Potter and Hodges & Taylor described the synthesis of 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92,** *Scheme 4.1*).[16,17] While Hodges and Taylor envisioned the easy preparation of 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzaldehyde (**88**, *Scheme 4.1*) by nitration of acetyl-5-chlorovanillin (**87**) with fuming nitric acid, acetyl-5-chlorovanillin (**87**) was not obtained and only the triacetylated derivative was isolated upon heating 5-chlorovanillin (**86**) in acetic anhydride at reflux. 5-Chlorovanillin (**86**) was next converted into the corresponding nitrile (**90**) by first forming the oxime, through reaction of the aldehyde **86** with hydroxylamine, and then the nitrile **90,** by oxime dehydration, before nitration was attempted a second time. A single product (**91**) was obtained in a high yield, but this nitro-derivative resisted hydrolysis and was therefore not useful in obtaining the target compound, 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**). No pure product could be isolated from the nitration of 4-acetoxy-5-chloro-3-methoxybenzoic acid (**93**), so Hodges and Taylor resorted to forming the methyl ester (**95**), through Fischer esterification of the carboxylic acid and acetylation of the phenol moiety, before attempting the nitration reaction again. This time, a mixture of two products was isolated, the major component being methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**). Base-mediated hydrolysis was then carried out successfully and the target compound **92** was obtained. We intended to form 6-chloro-2-methoxy-3-nitrophenol (**97**) through a subsequent decarboxylation reaction, which

could then undergo the key S_NAr reaction with commercially available 3-chloro-5-fluorobenzonitrile to produce 3-chloro-5-(6-chloro-2-methoxy-3-nitrophenoxy)benzonitrile (**99**, *Scheme 4.1*). Finally, reduction of the nitro group would be carried out to yield the final product, 3-(3-amino-6-chloro-2-methoxyphenoxy)-5-chlorobenzonitrile (**100**).

4.2.2 Nitrile formation

The unfortunate scientist, C. W. Scheele, who prepared hydrogen cyanide in 1782, died while attempting to isolate the anhydrous material. This event marked the first preparation of a nitrile, as well as, unsurprisingly, the start of nearly a century in which very little research was done in the field of nitrile chemistry. Nevertheless, a brave scientist, J. L. Gay-Lussac, succeeded in preparing the pure acid and establishing its composition in 1811 (without dying in the name of Science).[18] The first syntheses of nitriles were, however, only reported in 1832 by Wöhler and Liebig, who prepared benzoyl cyanide and benzonitrile, and in 1834 by Pelouze, who synthesised propionitrile.[18,19] The term “nitrile” was used for the first time by H. Fehling in 1844 when he named his newfound substance, synthesised by heating ammonium benzoate, benzonitrile.[18,20] Since the production of ammunition during World War I required large amounts of nitrate, attention was fixed on the Haber process of nitrogen fixation from air.[21] As a result of the increased availability of raw materials, the fifteen-year period from 1920 to 1935 showed a fourfold increase in the average number of papers published on nitrile chemistry per year.[18]

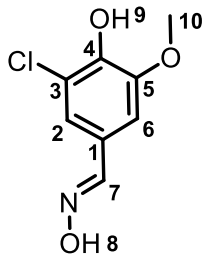
A number of methods have been developed for the preparation of nitriles, such as preparation by metathesis, cyanogenation of aromatic compounds by the Friedel-Crafts-Karrer synthesis or the Houben-Fischer synthesis, the addition of hydrogen cyanide to compounds containing various functional groups, amine dehydrogenation and the dehydration of oximes or amides, amongst others.[18] In order to convert the commercially available aldehyde, 3-chloro-4-hydroxy-5-methoxybenzaldehyde, into the aromatic nitrile **90**, Raiford and Potter suggested oxime dehydration, a reaction first described by Gabriel and Meyer in 1881 through the use of anhydride heated at reflux.[16,22] This is a reaction usually characterised by high yields achieved effortlessly using mild reaction conditions.[18]



*Scheme 4.1: The planned synthetic route to obtain the target diaryl ether **100**.*

4.2.2.1 Synthesis of 3-chloro-4-hydroxy-5-methoxybenzaldehyde oxime (89)[23]

Firstly, the aldehyde, 3-chloro-4-hydroxy-5-methoxybenzaldehyde (**86**), had to be converted into the corresponding oxime, to make dehydration to the nitrile possible. To this end, a method similar to the one described in Section 3.2.3.1, was used. Hydroxylamine hydrochloride together with pyridine was used to provide the nitrogen source, hydroxylamine, while activated molecular sieves were used to drive the equilibrium towards the desired product. After 3 hours of heating at reflux, followed by a work-up and product purification, a high yield of 92% of compound **89** was obtained.

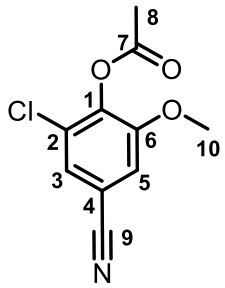


89

The success of the reaction was determined mainly through ^{1}H nuclear magnetic resonance (NMR) spectroscopic analysis. A singlet in the ^{1}H NMR spectrum at 11.05 ppm, provided evidence for N-OH (H_8). Furthermore, the $[M+H]^+$ ion was identified by high resolution mass spectrometry (HRMS) (calculated 202.0271, found 202.0273).

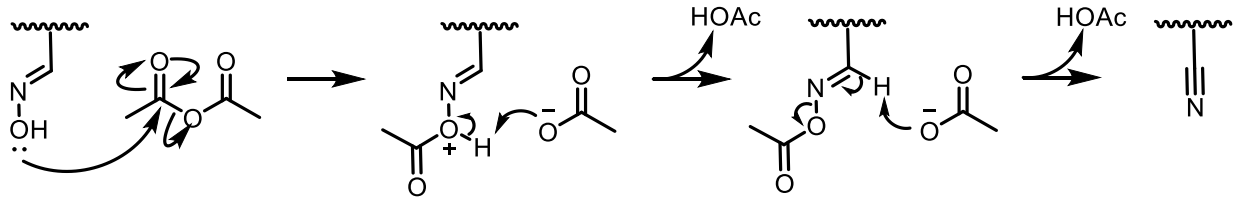
4.2.2.2 Synthesis of 2-chloro-4-cyano-6-methoxyphenyl acetate (90)[16,17]

The dehydration of the oxime **89** was subsequently achieved by heating in acetic anhydride at reflux. The reaction presumably commenced when the oxygen of the oxime (N-OH) attacked acetic anhydride at one of the C=O positions, resulting in the release of acetic acid and the formation of the *O*-acetyl oxime. A second molecule of acetic acid was produced when the HC-N hydrogen was removed to form the triple bond of the nitrile (*Scheme 4.2*). During this reaction, the phenol oxygen of compound **89** was also acetylated, as was expected. An 88% yield of compound **90** was obtained after work-up and purification by column chromatography.



90

The absence of the oxime N-OH signal in the ^{1}H NMR spectrum indicated nitrile formation. In addition, the acetylated phenol moiety was easily identified from the ^{1}H NMR spectrum when a singlet integrating for 3 protons was observed at 2.38 ppm. Furthermore, the ^{13}C NMR spectrum showed C_7 and C_8 at 167.2 ppm and 20.4 ppm respectively, characteristic of a C=O and a CH_3 group. The molecular ion was not confirmed by HRMS, but the correct compound structure was inferred from the confirmed structures that follow.



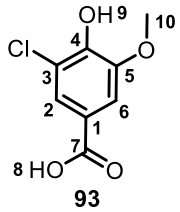
Scheme 4.2: Nitrile formation through oxime dehydration.

4.2.3 Selective nitration

Since Hodges and Taylor attained success from the nitration of methyl 4-acetoxy-3-chloro-5-methoxybenzoate (**95**), we set out to synthesise this compound by firstly hydrolysing the nitrile to the corresponding carboxylic acid, after which a Fischer esterification reaction, followed by acetylation of the phenol moiety, could yield product **95**. A possible nitrating agent, that could be used in a temperature controlled manner in the following step to produce methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**), was considered to be fuming nitric acid

4.2.3.1 Alkaline hydrolysis – Synthesis of 3-chloro-4-hydroxy-5-methoxybenzoic acid (93)[16,17]

The hydrolysis of nitriles has often been considered as one of the best and most trusted methods for the preparation of carboxylic acids.[24] It therefore comes as no surprise that this method was chosen by Hodges and Taylor. The reaction was achieved by heating 2-chloro-4-cyano-6-methoxyphenyl acetate (**90**), dissolved in a 6 M potassium hydroxide solution, at reflux. After acidification with concentrated hydrochloric acid to a pH of 3, a white precipitate formed and the pure title product **93** was obtained quantitatively after filtration. No further purification was required.



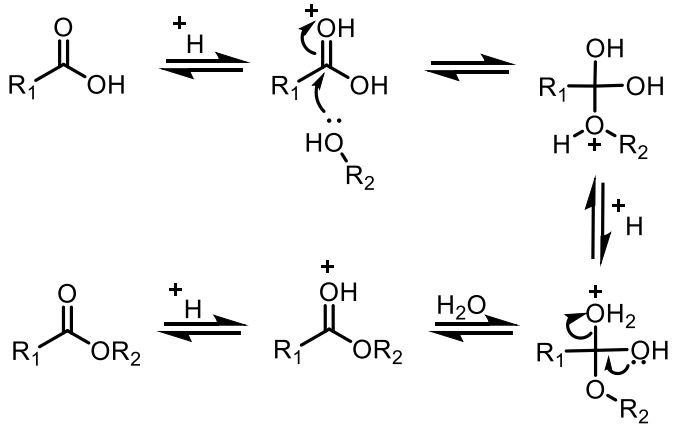
93

In terms of spectroscopic evaluation, the ^{1}H NMR spectroscopic signal of the phenol OH in the product was not detected. Rather, the lack of signals corresponding to the $OCOCH_3$ moiety, in the ^{1}H and ^{13}C NMR spectra, served as proof of deacetylation. With regards to the hydrolysis of the nitrile group, a C=O signal was detected at 166.3 ppm in the ^{13}C NMR spectrum. Conclusive evidence of product formation was provided by HRMS as the $[M-H]^-$ ion was identified and corresponded well to the expected mass.

4.2.3.2 Fischer esterification – Synthesis of methyl 3-chloro-4-hydroxy-5-methoxybenzoate (94)[17]

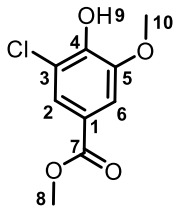
Fischer esterification can be explained generally as a reaction during which a carboxylic acid [$R_1C(O)OH$] is heated at reflux together with an alcohol (R_2OH), in the presence of an acid catalyst, to form the desired ester [$R_1C(O)OR_2$] (*Scheme 4.3*). Nucleophilic attack, by an alcohol molecule on the protonated carboxylic acid group, occurs first. The reaction then proceeds via an acyl-oxygen fission process, in which the bond between the original carbonyl-carbon atom and an oxygen of a hydroxyl group is cleaved.[25] This reaction was first described in 1895 in a paper published by Emil

Fischer and Arthur Speier and is sometimes referred to as the Fischer-Speier esterification reaction.[26]



Scheme 4.3: The mechanism of the Fischer esterification reaction.

In this case, the acid catalyst used was sulfuric acid. Since the methyl ester was desired, mainly because it is the simplest possible ester, methanol was used as solvent. 3-Chloro-4-hydroxy-5-methoxybenzoic acid (**93**), together with a catalytic amount of sulfuric acid dissolved in methanol, was thus heated at reflux for 3 hours. After work-up and purification by column chromatography, a good yield of 88% was obtained for the desired compound **94**.

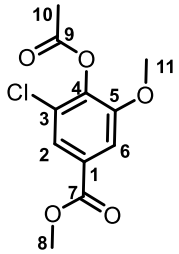


94

With respect to spectroscopic evaluation, a singlet integrating for 3 protons at 3.96 ppm in the ^{1}H NMR spectrum, representing H_8 of the product, confirmed ester formation. Once again, HRMS displayed the $[M-H]^-$ ion which corresponded well to the expected mass.

4.2.3.3 Acetylation – Synthesis of methyl 4-acetoxy-3-chloro-5-methoxybenzoate (95)[17]

Since the nitration of a phenol is difficult to control, because concentrated nitric acid may oxidize the phenol, protection by acetylation was proposed.[27] The acetylation of methyl 3-chloro-4-hydroxy-5-methoxybenzoate (**94**), using acetic anhydride at reflux, proceeded without difficulty and a 92% yield of compound **95** was obtained after work-up and product purification by column chromatography.



95

In terms of spectroscopic evaluation, the carbon of the OC=O moiety (C_9) presented as the most downfield signal in the ^{13}C NMR spectrum at 167.6 ppm, while C_{10} was observed as the furthest upfield signal at 20.4 ppm. In the ^{1}H NMR spectrum, H_{10} was seen as a large singlet, integrating for 3 protons, at 2.37 ppm. These results, together with the absence of the OH proton, present in the ^{1}H NMR spectrum of product **94**, lead to the conclusion that the acetylation of the OH group had been

successful. In addition, the molecular ion was confirmed by HRMS (calculated 259.0373, found 259.0374).

4.2.3.4 The key nitration reaction – Synthesis of methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (96)[17]

The nitration reaction, a general entry into aromatic nitrogen compounds, was first reported in 1834 by Mitscherlich who prepared nitrobenzene by treating benzene with fuming nitric acid. Since then, nitration has become a popular subject of study.[28] Since a nitro group withdraws electron density from the π-system, the ring becomes less reactive towards electrophiles. As a result, after the addition of the first nitro group, further nitration is difficult. In the unlikely event of a second nitration, this would occur *meta* to the first position.[27]

For this reaction to proceed, the powerful reagent, fuming nitric acid, was required. In order to ensure a maximum yield of the desired mono-substituted product, nitric acid was cooled to -30 °C before methyl 4-acetoxy-3-chloro-5-methoxybenzoate (**95**) was added in small portions, so as to keep the temperature between -30 and -20 °C. After addition, the reaction mixture was stirred for 1 hour at this temperature, before it was quenched on ice, neutralised and the organic products extracted. Only after extraction, could a thin layer chromatography (TLC) analysis of the reaction be performed. Three spots with different R_f values could be seen and these compounds were subsequently separated by column chromatography.

^{1}H NMR spectroscopy revealed that the three compounds were in fact the two possible mono-substituted products (**96** and **101**, *Figure 4.5*), as well as a small amount of the di-substituted compound (**102**). The challenge now was to determine which of the two mono-substituted products was the desired one. Various NMR spectroscopy experiments, such as 1D NOESY, 2D NOESY and 2D ROESYAD, using different mixing times and relaxation delays, were carried out, but no nuclear Overhauser effect (NOE) was observed when the protons indicated by yellow circles in *Figure 4.5* (**96**, **101**) were irradiated. We therefore resorted to our knowledge of chemical shifts and the factors that influence this. We determined that the aromatic proton-peak located furthest upfield, when comparing the spectra of **96** and **101**, should belong to product **101** due to the electron-shielding effects of the electron-donating methoxy group located next to this aromatic proton (*Figure 4.6*). This phenomenon was clearly seen from an overlay of the ^{1}H NMR spectra of the two mono-substituted products (*Figure 4.7*) and was supported by various predicted spectra of the two compounds. Deuterated chloroform was used as reference in both spectra before they were superimposed. In addition, the measured melting point of **96** (87 – 89 °C) compared well to that in the literature (89 – 90 °C) and is significantly lower than the literature melting point of **101** (140 – 141 °C).[17]

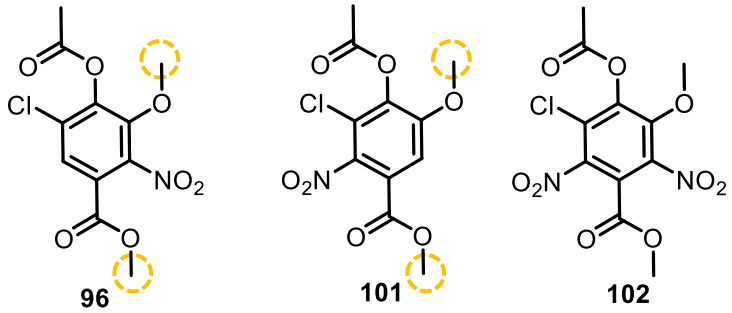


Figure 4.5: The protons irradiated for the1D NOESY, 2D NOESY and 2D ROESYAD NMR spectroscopy experiments.

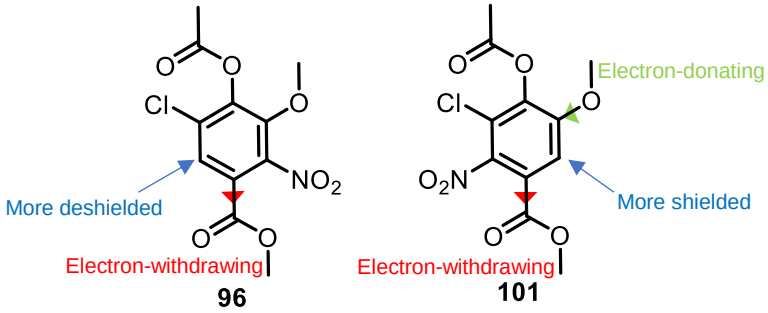


Figure 4.6: Electronic effects that could influence the position of the proton signals in ^{1}H NMR spectra of the two products.

A yield of 77% was obtained for product **96**, while the second mono-substituted product **101** was obtained in a 21% yield. The di-substituted product **102** was isolated in a 2% yield. Compound **96** was obtained in a greater yield of 86% when the time for which stirring was allowed, after addition of methyl 4-acetoxy-3-chloro-5-methoxybenzoate (**95**), was shortened to 30 minutes and the disubstituted product was no longer produced. These results were consistent with our expectations that **96** would be obtained as the major product, given the electron donating methoxy group, in contrast to the electron withdrawing chloro group.

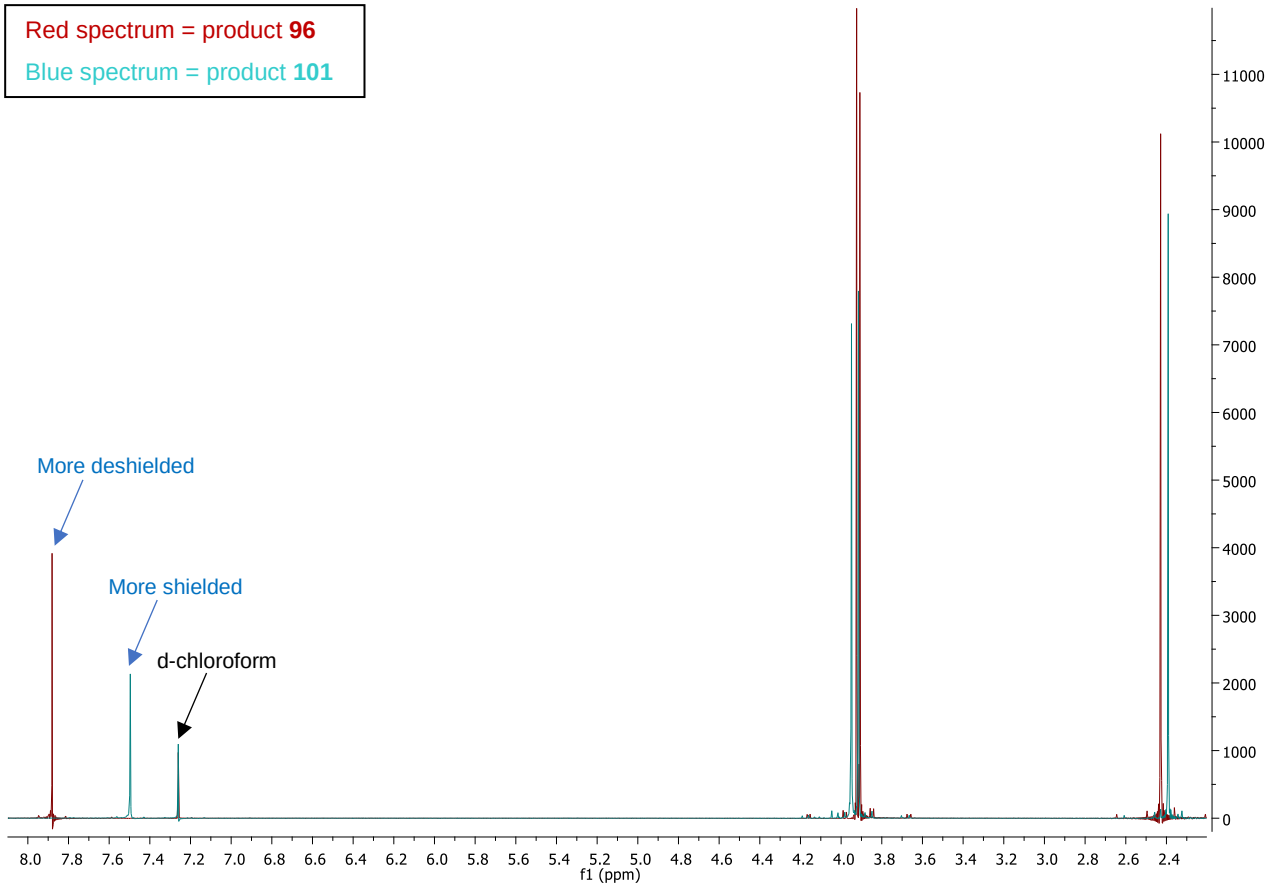
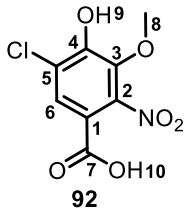


Figure 4.7: The superimposed ^{1}H NMR spectra of the two mono-substituted products from 8.1 ppm to 2.2 ppm.

4.2.4 Decarboxylation

4.2.4.1 Synthesis of 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (92)[17]

In order to obtain 6-chloro-2-methoxy-3-nitrophenol (**97**), the decarboxylation of 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**) was required. To this end, compound **92** was easily obtained when methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**) was heated at reflux in a methanol and sodium hydroxide solution. After work-up, the crude product was used as is in the next reaction.



92

From the ^{1}H NMR spectrum of product **92**, the absence of the two CH_3 groups, one belonging to the $OCOCH_3$ moiety and the other to the $COOCH_3$ moiety in compound **96**, confirmed product formation. A shift in the C_7 peak in the ^{13}C NMR spectrum of product **92** was also noted when compared to the spectrum of compound **96**. In addition, HRMS confirmed the presence of the molecular ion (calculated 245.9805, found 245.9816).

4.2.4.2 Synthesis of 6-chloro-2-methoxy-3-nitrophenol (97)

Decarboxylation involves the removal of a COOH group and replacing it with a hydrogen atom. The thermal decarboxylation of carboxylic acids is a well-known chemical transformation in organic

synthesis.[29] The usefulness of metals in this reaction, specifically copper, was only recognised in 1930 when Shepard *et al.* used stoichiometric amounts of copper or copper salts to decarboxylate furancarboxylic acids at high temperatures.[29] In general, heating and a metal catalyst are usually required.

The decarboxylation of 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**) proved to be a difficult task. Six different methods of decarboxylation were tried and are summarized below (*Table 4.5*). Each method is then discussed shortly.

*Table 4.5: Methods attempted to obtain the decarboxylated product **97**.*

Method	Metal source	Ligand	Additive	Solvent	Temperature (°C)	Time (h.)	Yield (%)
1	Cu_2O	1,10-phenanthroline	-	NMP:Quinoline 3:1	170	12	*a*
2	AgOAc	-	K_2CO_3	NMP	120	16	14%
3	-	-	-	Diphenyl ether	258	<0.5	-
4	Cu	-	-	Quinoline	238	4	-
5	Cu_2O	1,10-phenanthroline	-	NMP:Quinoline 3:1	170	1	48%

a = multiple spots

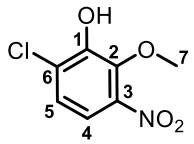
Yield = over two steps

4.2.4.2.1 Method 1

5-Chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**) was heated together with copper(I)oxide (Cu_2O) and 1,10-phenanthroline in a 3:1 mixture of *N*-methyl-2-pyrrolidone (NMP) and quinoline to 170 °C, with the aim of achieving decarboxylation to produce 6-chloro-2-methoxy-3-nitrophenol (**97**).[29] Under ultraviolet (UV) light, eight spots could be seen on the TLC plate. After work-up and purification by column chromatography, NMR spectroscopic analysis could unfortunately not be performed on the small quantities collected.

4.2.4.2.2 Method 2

Next, a reaction in which silver acetate (AgOAc) and potassium carbonate (K_2CO_3) were employed to decarboxylate compound **92**, was attempted. NMP was used as solvent and the reaction mixture was stirred at 120 °C.[30] After 16 hours, three spots were detected upon TLC analysis. After work-up and purification by column chromatography, one spot was identified as product from NMR spectroscopic analysis. A yield of only 14% (over two steps) was obtained, but the sample, later used for TLC analysis in future reactions, proved to be very valuable.



97

From the ^{1}H NMR spectrum, decarboxylation was confirmed by the observation of two doublets corresponding to the two aromatic protons at 7.43 ppm and 7.21 ppm. A large coupling constant of 9 Hz, denoting ortho coupling, additionally confirmed the correct positioning of the nitro group introduced in molecule **96**. A C=O signal was also no longer present in the ^{13}C NMR spectrum, indicative of the removal of this group. HRMS was used to confirm the presence of the $[M-H]^-$ ion which corresponded well to the expected mass.

4.2.4.2.3 Method 3

In search of a method that produced a higher yield of the product, diphenyl ether was heated to 260 °C and the carboxylic acid **92** was added portion-wise. After performing a work-up involving base extraction, it was hoped that compound **97** would be separated from diphenyl ether. However, upon TLC analysis, using the spotting sample acquired in method 2 as a reference for the desired product, it was seen that product **97** and diphenyl ether were both still present after work-up. Furthermore, the two compounds had the same R_f values making purification by column chromatography impossible. The very high boiling point of diphenyl ether also prohibited the use of distillation as a means of purification.

4.2.4.2.4 Method 4

Next, the carboxylic acid **92** was heated at reflux in quinoline, using copper as catalyst. After 4 hours, three spots could be seen on the TLC plate, none of which were product, as compared to the spotting sample obtained from method 2.

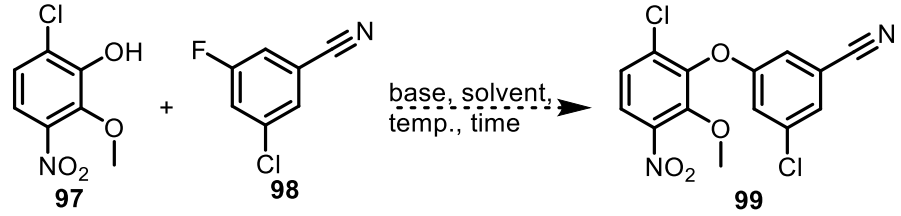
4.2.4.2.5 Method 5

Finally, a method with a reasonable yield of 48% (over two steps) of compound **97** was developed. When method 1 was repeated, but the reaction was carefully monitored by TLC analysis every 5 min., it was noted that the starting material was consumed after 1 h. When the reaction ran for longer than 1 h., additional spots started to appear on the TLC plates. It therefore seemed that the ideal reaction time was 1 h. and a maximum yield of 48% was obtained over two steps.

4.2.5 The attempted synthesis of 3-chloro-5-(6-chloro-2-methoxy-3-nitrophenoxy)benzonitrile (99)

The nucleophilic aromatic substitution (S_NAr) reaction mechanism involves the addition of a nucleophile, followed by the elimination of the leaving group. Commonly, an oxygen, nitrogen, or cyanide group acts as the nucleophile, while a halide is used as a leaving group. With a fluoro-compound known to react about 10^2 – 10^3 times faster than the chloro- or bromo-compounds and the iodo-compound reacting even slower, the 3-chloro-5-fluorobenzonitrile reagent (**98**), with F^- as leaving group, seemed ideal.[27] We therefore proceeded to try the reaction (*Scheme 4.4*).

At first, compound **97** was dissolved in dimethylformamide (DMF) together with compound **98** and the base, potassium carbonate (K_2CO_3), to pick up the extra proton on the OH after attack. The reaction was stirred at 80 °C for 18 hours, but only starting materials were detected on the TLC plate. The temperature was then increased to 100 °C and stirring was allowed to continue for a further 18 hours. At this point, when only starting materials were detected upon TLC analysis, a work-up was done and the organic compounds were purified using column chromatography, since it was possible that the product might have the same R_f as one of the starting materials. Unfortunately, from NMR spectroscopic analysis, this was not the case as only the two starting materials were isolated. We then attempted the reaction a second time using caesium carbonate (Cs_2CO_3) as base and acetonitrile (MeCN) as solvent.[31] This too resulted in the purification of starting materials after stirring at room temperature, 50 °C and at reflux. A final attempt to obtain the desired product made use of the higher temperature of 140 °C, K_2CO_3 and NMP.[32] In this case, decomposition seemed to be occurring. Three compounds were purified by column chromatography, but we were unsuccessful in determining their structures. *Table 4.6* summarizes the methods used and the results obtained.



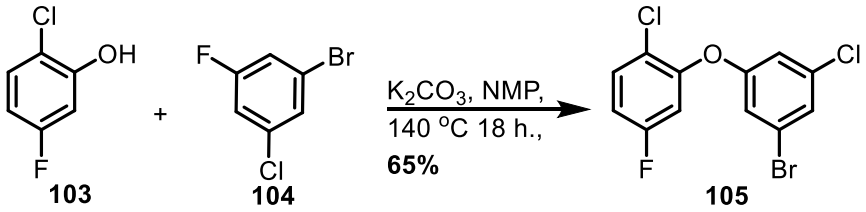
Scheme 4.4: A nucleophilic aromatic substitution reaction attempted under the conditions in Table 4.6.

*Table 4.6: Synthetic methods attempted with the hope of synthesising **99**.*

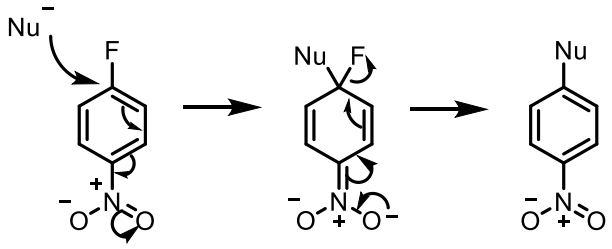
#	Base	Solvent	Temperature (°C)	Time (h.)	Result
1	K_2CO_3	DMF	80	18	SM
2	K_2CO_3	DMF	100	18	SM
3	Cs_2CO_3	MeCN	RT	2	SM
4	Cs_2CO_3	MeCN	50	20	SM
5	Cs_2CO_3	MeCN	Reflux	20	SM
6	K_2CO_3	NMP	140	40	3 unidentified compounds

SM = starting materials

Although proof of this type of reaction, specifically between 2-chloro-5-fluorophenol (**103**) and 1-bromo-3-chloro-5-fluorobenzene (**104**) to form 2-(3-bromo-5-chlorophenoxy)-1-chloro-4-fluorobenzene (**105**) (*Scheme 4.5*), was found in good yield (65%) in the literature[32], an electron-withdrawing group, such as a carbonyl, nitro, or cyanide group, is usually required in an *ortho*- or *para*-position to the leaving group (*Scheme 4.6*).[27] This is important for electronic stabilization upon attack by the nucleophile.[27]



Scheme 4.5: An example of a similar S_NAr reaction by Gomez et al.[32]



Scheme 4.6: The mechanism by which electrons flow out of the ring and into and anion-stabilizing group during S_NAr.

4.3 Proof of concept compounds

In light of the success accomplished by Gomez *et al.*, a number of question about our scaffold arose. Since the nucleophile used by Gomez *et al.* had a single *ortho* position occupied, we wondered if steric hindrance may play a role in the reaction. Secondly, we questioned whether the nitro group in the nucleophile is too electron-withdrawing, making the lone pair on the OH less readily available for attack. We also wondered if the problems could be avoided by using an Ullmann condensation reaction. To this end, a number of reactions were carried out in an attempt to answer these questions.

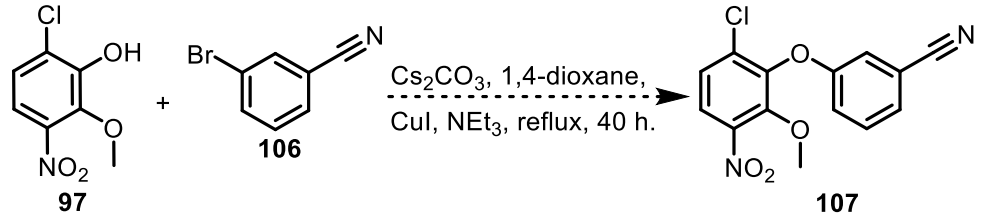
4.3.1 Ullmann condensation

The Ullmann condensation reaction, first described by Fritz Ullmann in 1905, is the copper-mediated arylation of amines, phenols, amides, carbamates and activated methylene compounds.[33,34] Interestingly, the formation of a nitrogen-carbon bond to an aromatic halide in the presence of copper, with the nitrogen-containing component being an amide, is the only named reaction for which a woman is unambiguously recognised. This reaction, called the Goldberg reaction, was named after Irma Goldberg, Fritz Ullmann's wife.[35]

These coupling reactions have numerous industrial applications, such as the synthesis of intermediates and synthetic targets for the life sciences and polymer industries[34] and was attempted to overcome the problem experienced in this part of the project.

4.3.1.1 The attempted synthesis of 3-(6-chloro-2-methoxy-3-nitrophenoxy) benzonitrile (107)

With 3-bromobenzonitrile (**106**) available in our laboratories, a test reaction could be carried out using this aromatic halide, 6-chloro-2-methoxy-3-nitrophenol (**97**), copper(I)iodide (CuI), Cs_2CO_3 and triethylamine (NEt_3) heated at reflux in 1,4-dioxane (*Scheme 4.7*). After 40 hours, no reaction had occurred and the starting materials were recovered following purification by column chromatography. Ullmann condensation was therefore eliminated as a possible solution.



*Scheme 4.7: The attempted Ullmann condensation of 6-chloro-2-methoxy-3-nitrophenol (**97**) and 3-bromobenzonitrile (**106**).*

4.3.2 An investigation on steric hinderance and the effects of the nitro group

In order to eliminate the effects of the nitro group, while investigating the sole role of the two groups *ortho* to the reacting OH, the synthesis of 3-chloro-5-(2-chloro-6-methoxyphenoxy)benzonitrile (**108**, *Figure 4.8*) was attempted. Next, steric effects were eliminated, while the nitro group was left in place through the synthesis of 3-chloro-5-(3-nitrophenoxy)benzonitrile (**109**).

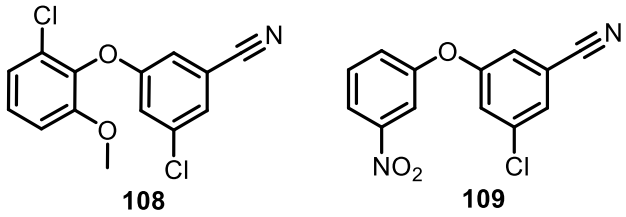
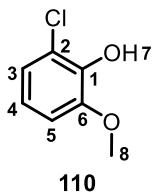


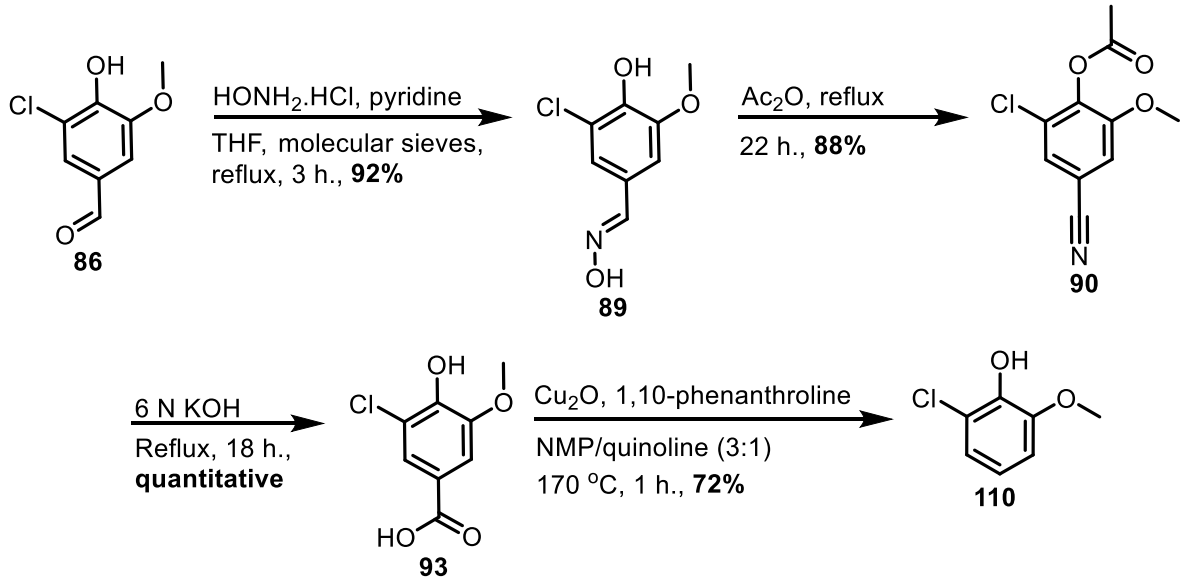
Figure 4.8: Target compounds for an investigation on steric and electronic effects playing a role during the S_NAr reaction.

4.3.2.1 Decarboxylation – Synthesis of 2-chloro-6-methoxyphenol (110)

The same method as per the synthesis of compound **97** (Section 4.2.4.2) was used (see *Scheme 4.8*). Product formation was achieved in good yield (72%) by the addition of Cu_2O and 1,10-phenanthroline to a solution of 3-chloro-4-hydroxy-5-methoxybenzoic acid (**93**) in NMP and quinoline.



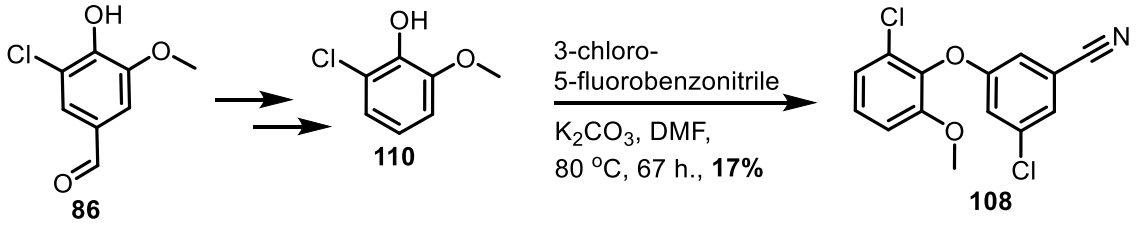
The three aromatic protons, as well as the three aliphatic OCH_3 protons, were accounted for by ^1H NMR spectroscopic analysis of product **110**. A clear loss of the COOH signal in the ^{13}C NMR spectrum, confirmed product formation. Additionally, the NMR spectra of the product corresponded well to those reported in the literature.[36]



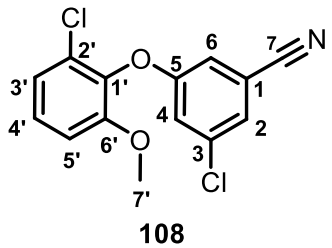
*Scheme 4.8: The synthetic route to compound **110**, for use in the S_NAr reaction.*

4.3.2.2 Synthesis of 3-chloro-5-(2-chloro-6-methoxyphenoxy)benzonitrile (108)

Next, compound **110**, 3-chloro-5-fluorobenzonitrile (**98**) and K_2CO_3 were added to a Schenk tube, heated to 80 °C in DMF and stirred at this temperature for 18 hours. TLC analysis showed the formation of a new spot, but both starting materials were also still present. The reaction mixture was stirred for a total of 67 hours, during which frequent TLC plates were run. When no significant change was noted, a work-up was done and the new spot was purified by column chromatography. NMR spectroscopy analysis revealed that our first diaryl ether product **108** had been synthesized in a poor 17% yield (*Scheme 4.9*).



*Scheme 4.9: The synthesis of our first diaryl ether compound **108**.*



The furthest downfield peak, located at 7.30 ppm, in the aromatic region of the ^{1}H NMR spectrum of product **108** (shown in *Figure 4.9*), was identified as one of the three protons on the 3-chloro-5-benzonitrile half of the molecule based on the small *J* coupling values of $J_1 = J_2 = 1.4$ Hz. The doublet of doublets, at 7.22 ppm, with the larger J coupling values of $J_1 = J_2 = 8.3$ Hz, thus belonged to $H_{4'}$. Together with the two multiplets at 7.12-7.09 ppm and 7.01-6.88 ppm, each integrating for 2 protons, a total of 6 aromatic protons were accounted for. The protons labeled $H_{7'}$ was seen as a singlet at 3.80 ppm, integrating for 3 protons. HRMS was used to confirm the formation of the product and the $[M+H]^+$ ion was detected (calculated 294.0089, found 294.0079).

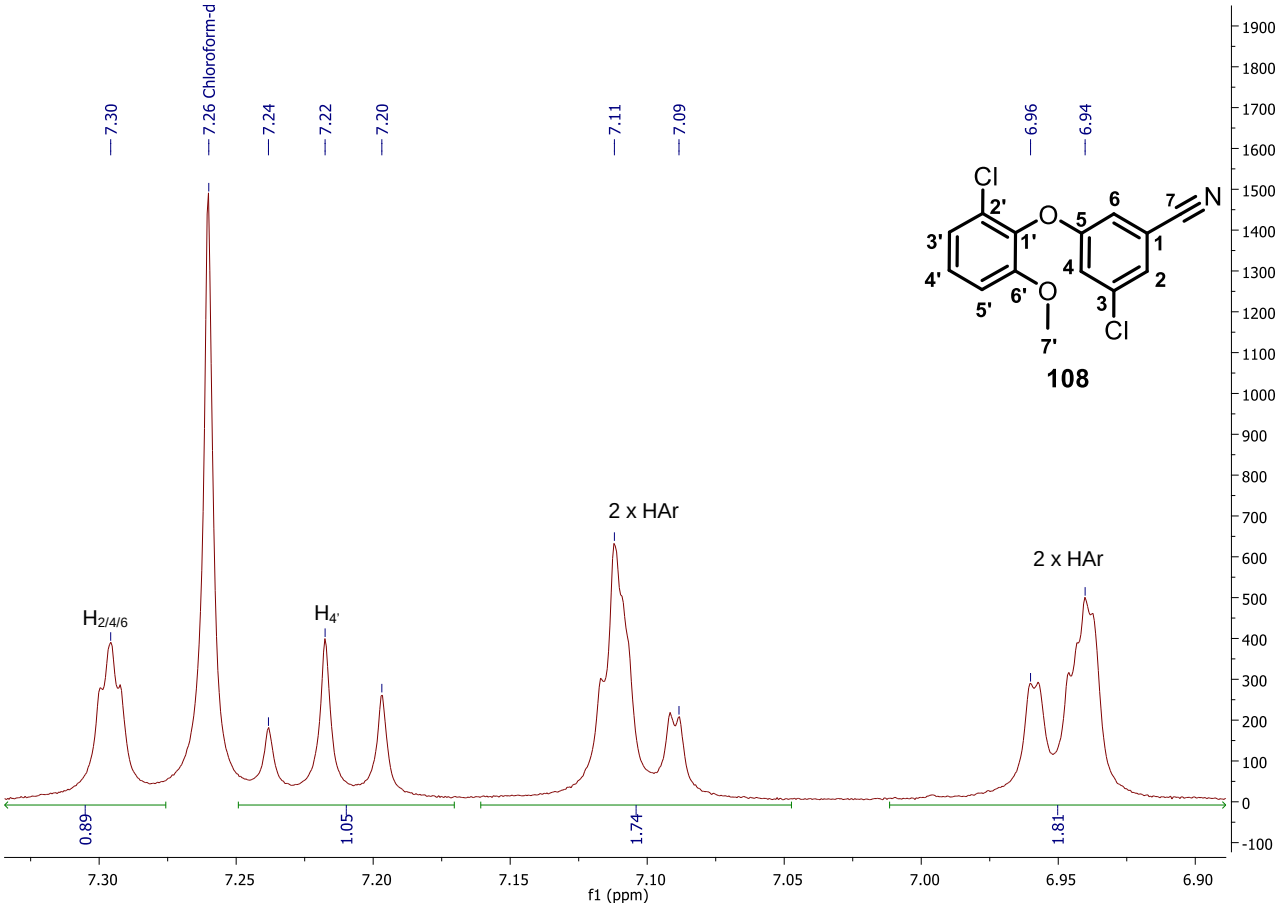
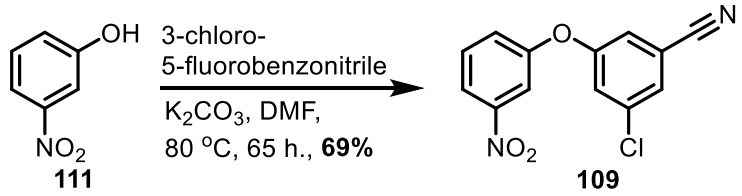


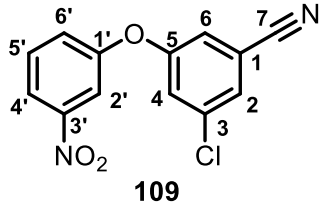
*Figure 4.9: A ^{1}H NMR spectrum of **108** showing the aromatic region of product.*

4.3.2.3 Synthesis of 3-chloro-5-(3-nitrophenoxy)benzonitrile (109)

As per the synthesis of compound **108**, the two aromatic compounds, in this case 3-nitrophenol (**111**) and 3-chloro-5-fluorobenzonitrile (**98**), and K_2CO_3 were dissolved in DMF and stirred at 80 °C for 65 hours (*Scheme 4.10*). The title product was obtained in a surprisingly good yield of 69% after work-up and column chromatography.



*Scheme 4.10: A second successful S_NAr reaction between 3-nitrophenol (**111**) and 3-chloro-5-fluorobenzonitrile (**98**) to produce diaryl ether **109**.*



In terms of spectroscopic evaluation, the ^{1}H NMR spectrum showed all aromatic protons clearly (*Figure 4.10*). The doublet of doublets of doublets at 8.10 ppm was identified as $H_{4'}$, with the *J*-coupling values matching those of $H_{5'}$, a doublet of doublets at 7.61 ppm, $H_{6'}$, a doublet of doublets of doublets at 7.39 ppm, and $H_{2'}$, a doublet of doublets at 7.18

ppm. The three doublets of doublets, H_6, H_2 and H_4, were observed at 7.88 ppm, 7.45 ppm and 7.26 ppm.

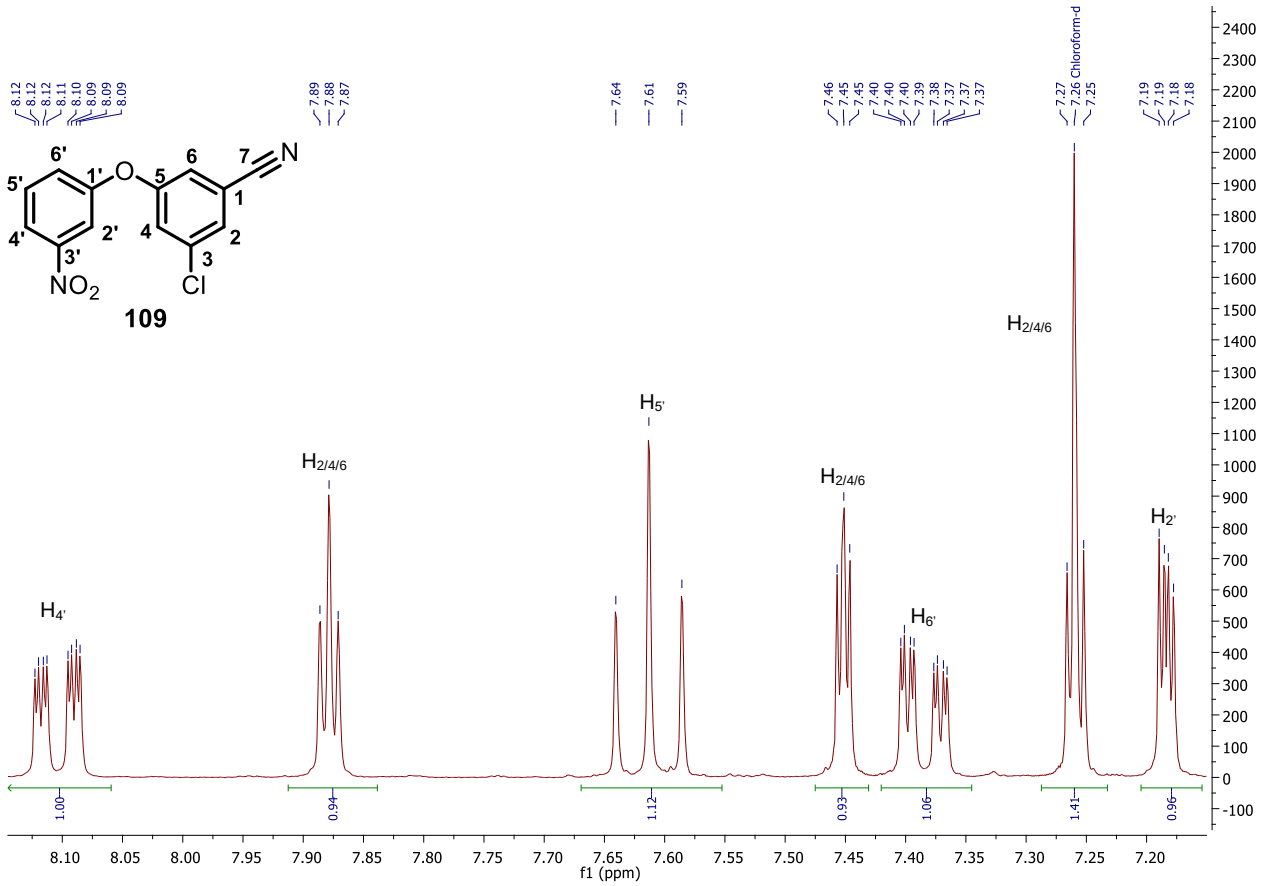


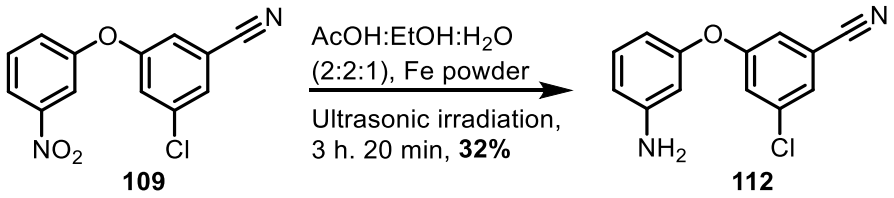
*Figure 4.10: The ^{1}H NMR spectrum of **109** showing only the relevant aromatic region.*

4.3.2.4 Aryl nitro reduction – synthesis of 3-(3-aminophenoxy)-5-chlorobenzonitrile (112)

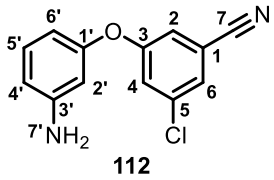
Aromatic amines are common intermediates in the synthesis of dyes, antioxidants and photographic, pharmaceutical and agricultural chemicals. These compounds may be obtained by the reduction of aromatic nitro compounds and, as a result, this chemical reaction has gained much attention.[37] Activated metal catalysis, such as zinc, tin, iron or Raney nickel, and transition-metal-catalyzed hydrogenation are commonly used to achieve the reduction.[37–41] The downfall of these methods are the often harsh reaction conditions required and the lack of selectivity in the presence of reduction-sensitive functional groups.[41] Selective reduction of aryl nitro compounds have, however, been reported with the use of iron powder and dilute acid, a method first developed in 1930 by Jenkins *et al.*[42–45] Common disadvantages of this method include high reaction temperatures, long reaction times, possible halogenation reactions and the unsuitability of use with acid-sensitive functional groups.[43] At high temperatures, the use of acetic acid in the reduction of aromatic nitro groups, in the presence acid sensitive functionalities, have been reported successful.[46] Drawing inspiration from these positive results, Gamble *et al.* reasoned that ultrasonication, known to increase the catalytic activity of metal particles by factors as high as 10^5,[47] may be used together with iron

activated by acetic acid to improve reaction times, lower reaction temperatures and increase the yields of aromatic nitro reduction reactions in the presence of sensitive functionalities. Finally, a selective reduction method of aryl nitro compounds had been developed.[43]

This method was used to reduce the aryl nitro group in compound **109** in the presence of its chloro- and nitrile-moieties (*Scheme 4.11*). To achieve this selective reduction, 3-chloro-5-(3-nitrophenoxy)benzonitrile (**109**) was dissolved in a mixture of glacial acetic acid, ethanol and water. Reduced iron powder was then added and the suspension was exposed to ultrasonic irradiation. After 3 hours and 20 minutes, the reaction mixture was filtered through Celite and a 32% yield of the title product was obtained after work-up and purification by column chromatography. An improved yield may be obtained with a longer reaction time, since unreacted starting material was isolated as the only other organic component of the reaction mixture.



*Scheme 4.11: A selective nitro reduction reaction to yield final compound **112** from diaryl ether **109**.*



An obvious indication of product formation was the singlet corresponding to the two NH_2 protons in the ^1H NMR spectrum of the product. This peak was located at 3.80 ppm and integrated for 2 protons, as expected. Furthermore, HRMS confirmed product formation with the identification of the $[M+H]^+$ ion (calculated 245.0482, found 245.0484).

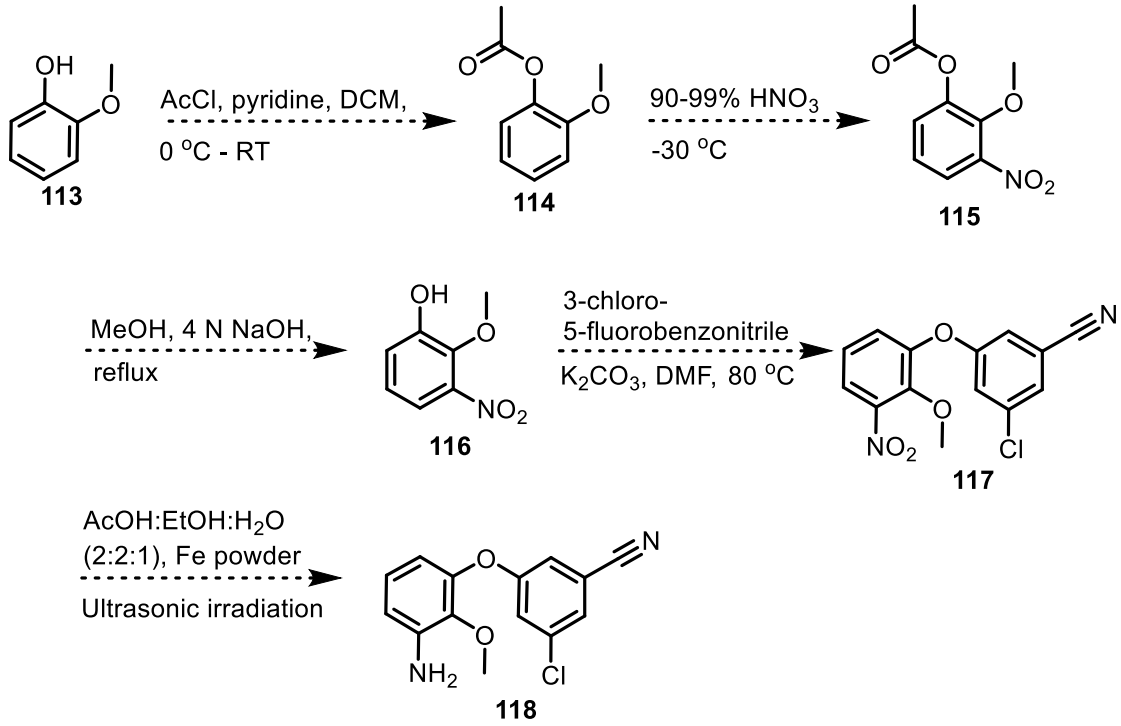
4.3.2.5 Conclusions that can be drawn

From the above short study on the steric effects of the chloro and methoxy moieties and the electronic effects of the nitro group, it can be concluded that, separately, these effects do not prevent the formation of the desired product 3-chloro-5-(6-chloro-2-methoxy-3-nitrophenoxy)benzonitrile (**99**). It is, however, possible that, when combined, the steric and electronic effects result in an energy barrier too great to overcome for product formation.

4.3.3 Second best

Since the synthesis of the desired final product, 3-(3-amino-6-chloro-2-methoxyphenoxy)-5-chlorobenzonitrile (**100**), was not successful, an alternative diaryl ether, consisting of the most important features of compound **100,** was designed by means of molecular modelling. The amine group, forming a hydrogen bonding interaction with Lys101, and the methoxy group, occupying the Val179 pocket, were hypothesized to be the key functional groups important for activity. The chloro-moiety, of seemingly less importance, was omitted. In terms of the synthesis, a very similar strategy

to that implemented for the synthesis of compounds **100**, **108** and **109** was designed to synthesize the final target, 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile (**118**), in this section. We envisaged the successful selective nitration of the commercially available 2-methoxyphenol (**113**), commonly known as guaiacol, by first protecting the phenol OH, before treating the resultant compound with fuming nitric acid at –30 °C to form compound **115** (*Scheme 4.12*). An easy deprotection was anticipated with sodium hydroxide heated at reflux, after which the S_NAr reaction would be attempted, as was done in the previous cases, to form product **117**. The selective nitro reduction could then be achieved through the use of activated iron powder and ultrasonic irradiation to result in the desired final product **118**.

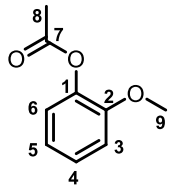


*Scheme 4.12: The planned synthetic route to compound **118**, 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile.*

4.3.3.1 Selective nitration

4.3.3.1.1 Synthesis of 2-methoxyphenyl acetate (114)

In order to obtain the *O*-acetylguaiacol derivative **114**, acetyl chloride was added dropwise to a solution of guaiacol (**113**) and pyridine in dichloromethane (DCM) at 0 °C. After addition, the solution was allowed to warm to room temperature, while stirring was continued for 22 hours. A long reaction time was allowed for complete conversion, since the starting material and product had exactly the same R_f. This highly effective protection reaction resulted in a 94% yield of the desired product, after work-up and product purification.

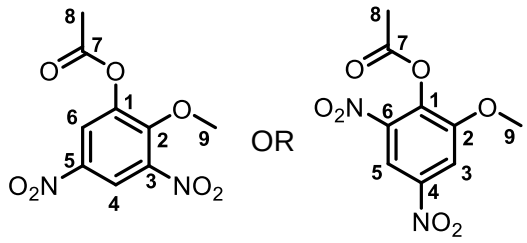


114

In terms of structural confirmation, the CH_3 protons of the acetate group that had been added were clearly visible as a singlet, integrating for 3 protons at 2.32 ppm, in the ^1H NMR spectrum of the product. In addition, C=O was observed as the furthest downfield peak at 169.2 ppm in the ^{13}C NMR spectrum. It should be noted that the NMR spectra compares well to those reported in the literature.[48]

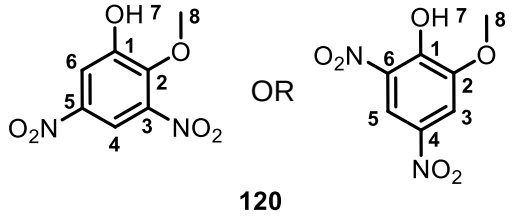
4.3.3.1.2 Attempted synthesis of 2-methoxy-3-nitrophenyl acetate (115)

As with the synthesis of methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**), fuming nitric acid was cooled to –30 °C before 2-methoxyphenyl acetate (**114**) was added portion-wise. After addition, the solution was stirred for a further 30 minutes before a work-up was done. Two products, mostly isolated as a mixture, were obtained after purification by column chromatography. From NMR spectroscopic analysis, it was clear that in both products, double nitration had occurred, since two doublets, instead of the expected three doublet of doublets, were seen. The first spot on the TLC plate, with an R_f of 0.63 in a 50% EtOAc/Hexane solution, was a dinitro acetate compound (**119**), while the compound corresponding to the second spot, with an R_f of 0.53 (50% EtOAc/Hexane), had undergone deacetylation (**120**). We were unable to determine the nitration positions through 2D NMR spectroscopy (NOESY, CIGARD and COSY), but the relatively small *J* coupling values indicated 3,5- or 4,6-dinitro substitution.



119

The ^1H NMR spectrum of the two possible acetylated products alongside showed two aliphatic singlets at 2.43 ppm and at 4.05 ppm, representing H_8 and H_9 respectively. The two aromatic doublets at 8.23 ppm and 8.60 ppm, with a small *J* coupling value of 2.8 Hz, indicated nitration in either the 3 and 5 positions or the 4 and 6 positions.



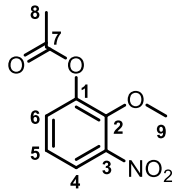
120

The clear absence of two carbon peaks corresponding to the acetate group in the ^{13}C NMR spectrum of the second product, indicated deacetylation. Again, from the small *J* coupling value of the two doublets (2.7 Hz), the locations of nitration were deduced to be meta to one another.

4.3.3.1.3 Synthesis of 2-methoxy-3-nitrophenyl acetate (115)

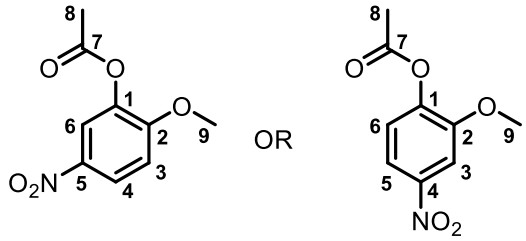
An alternative method towards the synthesis of compound **115** was then attempted. Guaiacol acetate (**114**) was first dissolved in acetic anhydride, after which ice cooled fuming nitric acid, also dissolved in acetic anhydride, was added dropwise. Directly after addition, a work-up was done. After purification by column chromatography, the desired product was obtained in an acceptable 73% yield. The by-product, either 2-methoxy-4-nitrophenyl acetate or 2-methoxy-5-nitrophenyl acetate

(**121**), was obtained in a 27% yield. The results were consistent with our expectations that compound **115** would be obtained as the major product due to electronic effects of the adjacent methoxy group.



115

The ^{1}H NMR spectrum of the product **115** showed three doublet of doublets in the aromatic region, two of which had *J* coupling values of 8.1 Hz and 1.7 Hz and the third with $J_1 = J_2 = 8.1$ Hz. These coupling constants places the nitro moiety in either the 3 or the 6 position. A definitive answer was obtained from a literature comparison of the ^{1}H NMR spectra of the deacetylated product (see **116**, Section 4.3.3.1.4).

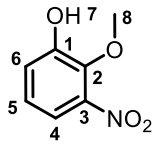


121

For the by-product **121**, the coupling constants of the one doublet of doublets and the two doublets in the ^{1}H NMR spectrum of the by-product, could only result from nitration in two possible positions namely the 4 or the 5 positions. The doublet of doublets (*J* = 9.1, 2.7 Hz) resulted from coupling with one vicinal proton and a proton in a *meta* relationship with it. The doublet with the large *J* coupling constant of 9.1 Hz resulted from coupling to one vicinal proton, while the doublet with the smaller *J* coupling constant of 2.7 Hz resulted from coupling to a *meta* proton.

4.3.3.1.4 Synthesis of 2-methoxy-3-nitrophenol (116)

As anticipated, the deacetylation, achieved by heating compound **115** in a methanol and NaOH solution at reflux, proceeded without difficulty and, after purification, the title product was obtained in 91% yield.



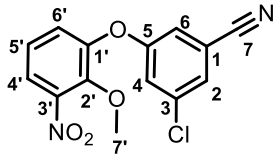
116

The NMR spectra for **116** showed the loss of the acetate moiety, while a singlet at 5.99 ppm in the ^{1}H NMR spectrum represents the OH proton. Additionally, the spectra compared well to those reported in the literature, confirming the correct positioning of the nitro group (refer to product **115**, Section 4.3.3.1.3).[49]

4.3.3.2 The third diaryl ether compound

4.3.3.2.1 Synthesis of 3-chloro-5-(2-methoxy-3-nitrophenoxy)benzonitrile (117)

The nucleophilic aromatic substitution reaction proceeded when 2-methoxy-3-nitrophenol (**116**) and 3-chloro-5-fluorobenzonitrile (**98**) were dissolved in DMF, together with K_2CO_3, and the reaction mixture was heated to 80°C. After 48 hours, followed by a work-up and product purification, the title product was obtained in a low 23% yield.

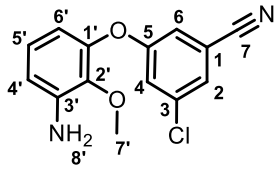


117

In terms of spectroscopic evaluation, all aromatic proton atoms were accounted for from the ^{1}H NMR spectrum of the product **117**. In addition, the aliphatic CH_3 signal was observed as a large singlet at 3.95 ppm. All carbon atoms were also accounted for from the ^{13}C NMR spectrum of the product.

4.3.3.2.2 Synthesis of 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile (118)

The only remaining reaction, to complete the synthesis, was the reduction of the nitro group. To this end, 3-chloro-5-(2-methoxy-3-nitrophenoxy)benzonitrile (**117**) and reduced iron powder were added to a mixture of glacial acetic acid, ethanol and water and the suspension was exposed to ultrasonic irradiation. After 3 hours and 40 minutes, the reaction mixture was filtered through Celite and the product was purified using column chromatography, resulting in a 79% yield of compound **118**.



118

All proton and carbon atoms were accounted for from the ^{1}H and ^{13}C NMR spectra. The ^{1}H NMR spectrum below clearly shows the broad singlet at 3.97 ppm corresponding to the two NH_2 protons (*Figure 4.11*). This confirmed a successful nitro reduction reaction. Furthermore, the $[M+H]^+$ ion was identified by HRMS (calculated 275.0587, found 275.0582).

Though in a poor overall yield of 11%, the synthesis of 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile (**118**) was successful (full scheme shown in *Scheme 4.13*). In retrospect, the incorporation of an electron-withdrawing group, *ortho* or *para* to the leaving group, may have resulted in better yields in the S_NAr reactions.[27]

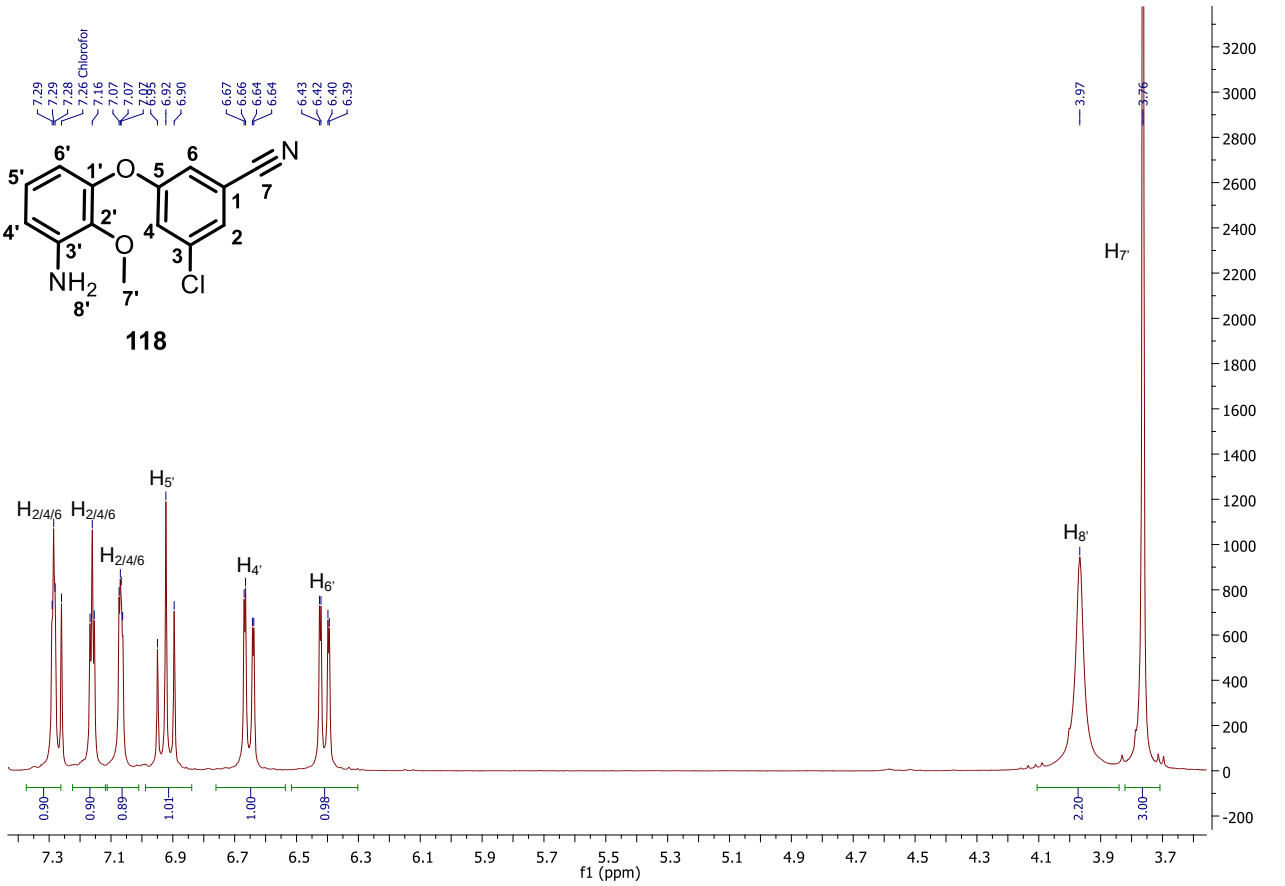
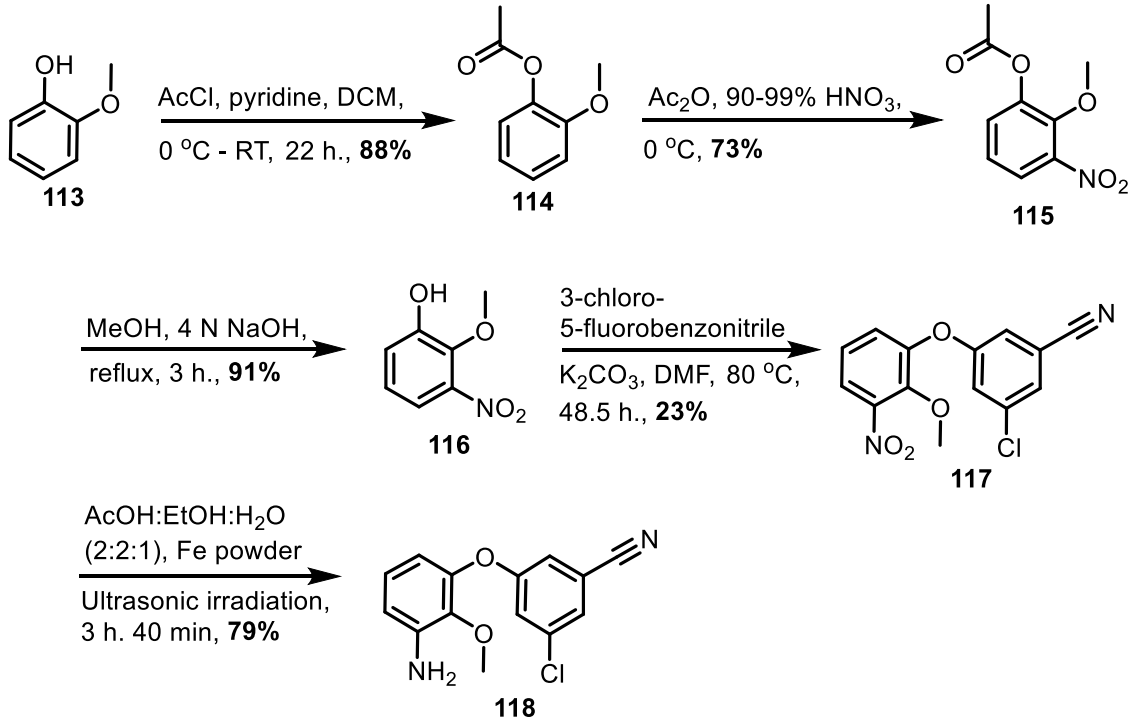


*Figure 4.11: The ^{1}H NMR spectrum of **118** showing all aromatic protons, the two NH_2 protons and the OCH_3 group.*



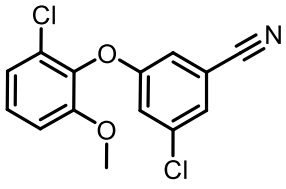
*Scheme 4.13: The complete synthesis of 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile (**118**).*

4.4 Biological results and concluding remarks

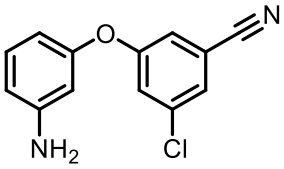
The three proof of concept compounds **108**, **112** and **118** were sent to our collaborator, Dr. Adriaan Basson at the University of the Witwatersrand, for biological evaluation. An *in vitro* single-cycle, non-replicative phenotypic assay, using a HIV-1 retroviral vector system for the production of virus-like particles, was used.[50] The compounds were incubated with 293T cells together with the virus-like particles and after 48 hours, inhibition was quantified by luminescence measurement. The toxicity and activity assay results are presented in *Table 4.7* below. Note that nevirapine, used as control, had an IC_{50} value within the expected range.

From the efficacy results of the three diaryl ether compounds, **118** proved to be highly active with an IC_{50} value of 5 nM (*Table 4.7*, highlighted in peach). When the methoxy and amine moieties were introduced on two separate scaffolds, compounds **108** and **112** respectively, poorly active compounds resulted. When these groups were combined, a significant improvement in activity was noted. Future studies will thus focus on incorporating the chloro group onto the same scaffold with the aim of synthesizing a superior NNRTI (see Future work, Chapter 6, Section 6.3). The high toxicity values of these compounds are additionally encouraging since the high values are synonymous to low toxicity. This implies that a large quantity of the compounds will be needed before resulting in a toxic effect.

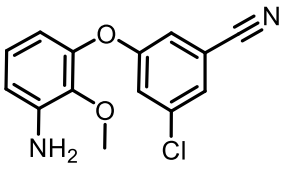
Table 4.7: The toxicity and activity results for the three proof of concept compounds.



108



112



118

Compounds	Toxicity (CC_{50}, μM)				Activity (IC_{50}, μM)			
	1	2	Ave	SD	1	2	Ave	SD
108	>100	>100	>100	-	0.202	0.218	0.210	0.011
112	>100	>100	>100	-	1.051	1.193	1.122	0.100
118	98.03	94.64	96.33	2.40	0.006	0.004	0.005	0.001
NVP ctrl.					0.105	0.076	0.091	0.20

Ave = average

SD = standard deviation

NVP ctrl. = Nevirapine control

Peach highlight = the most potent compound against wild-type HIV-1.

Compound **118** was further subject to testing against a panel of mutant strains (*Table 4.8*). Unfortunately, this compound was found to be ineffective against most of the prevalent NNRTI resistance mutations. Nevertheless, activity was retained against the important Y188C strain, which renders NVP ineffective. This validates further optimization of this scaffold toward a diaryl ether with an improved resistance profile.

Table 4.8: A summary of the activity results of compound 118 against a panel of mutant HIV viral strains.

118	Activity (IC_{50}, μM)							
	WT	K103N	Y181C	K103N+Y181C	V106M	Y188C	Y188H	G190A
Ave	0.007	0.173	0.171	0.957	0.271	0.029	0.508	0.398
SD	0.004	0.035	0.008	0.016	0.012	0.006	0.130	0.116
FC	1.0	23.9	23.6	132.5	37.5	4.0	70.3	55.2

Ave = average

SD = standard deviation

FC = Fold change in IC_{50} relative to the wild-type

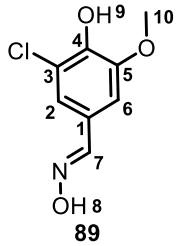
Yellow highlight = retained activity against the Y188C mutant viral strain

4.5 Experimental section

Please refer to Section 2.4.1, Chapter 2 for general procedures pertaining to the purification of reagents and solvents, chromatography, spectroscopic and physical data, molecular modelling and other general procedures followed. It should be noted that an arbitrary numbering system was used and that reference should be made to the provided structure when studying the NMR spectra.

4.5.1 Synthesis of 3-chloro-4-hydroxy-5-methoxybenzaldehyde oxime (89)

To a 50 mL two-neck round-bottom flask fitted with a condenser, which had been flushed with alternating vacuum and nitrogen purge cycles and was charged with dry THF (20 mL), approximately 15 activated molecular sieves were added, followed by 3-chloro-4-hydroxy-5-methoxybenzaldehyde (1.00 g, 5.36 mmol), hydroxylamine hydrochloride (2.5 eq., 0.931 g, 13.4 mmol) and dry pyridine (2 eq., 0.90 mL, 11 mmol). The reaction mixture was then heated at reflux for 3 h., after which it was cooled to RT and diluted with H_2O (40 mL). $NaHCO_3$ was then added until a pH of 7 was obtained. The product was extracted with Et_2O (3 × 50 mL), the combined organic layers were dried over $MgSO_4$ and the solvent was evaporated. Purification was done by column chromatography (5 – 50% EtOAc/Hexane) to yield the title compound (0.990 g, 4.91 mmol, 92%) (R_f = 0.39, 50% EtOAc/Hexane) as a light-yellow solid.

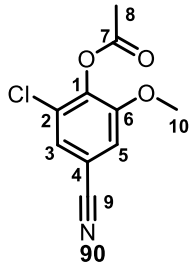


89

Decomposed at 160 °C. **IR (ATR, cm^{-1})** 3455, 3400 – 2600 (O-H str), 1679 (C-N str), 1603 and 1509 (C=C str), 1419, 1291 (C-O str), 1173 (C-O str), 1039, 960 (C-H out-of-plane bend), 851 (C-Cl str). **^{1}H NMR (400 MHz, DMSO-d) δ** 11.05 (s, 1H, H_8), 9.75 (s, 1H, H_9), 8.01 (s, 1H, H_7), 7.19 – 7.12 (m, 2H, H_2 and H_6), 3.84 (s, 3H, H_{10}). **^{13}C NMR (101 MHz, DMSO-d) δ** 148.8, 147.1, 143.9, 124.7, 120.4, 119.9, 107.5, 56.1 (C_{10}). **HRMS** calculated for $C_8H_9NO_3Cl$ $[M+H]^+$, 202.0271, found 202.0273.

4.5.2 Synthesis of 2-chloro-4-cyano-6-methoxyphenyl acetate (90)

3-Chloro-4-hydroxy-5-methoxybenzaldehyde oxime (**89**) (0.893 g, 4.43 mmol) in acetic anhydride (Ac_2O) (10 mL) was heated at reflux in a two-neck round-bottom flask for 18 h. The reaction mixture was then poured into a beaker containing several volumes of ice-water and was stirred. The mixture was subsequently neutralized with a potassium hydroxide solution (1 M) and the product was extracted with EtOAc (3 x 30 mL). The combined organic layers were then dried over $MgSO_4$, filtered and the solvent was removed under reduced pressure. Purification by column chromatography (5 – 50% EtOAc/Hexane) yielded the title compound (0.833 g, 3.69 mmol, 88%) (R_f = 0.75, 50% EtOAc/Hexane) as a white solid.

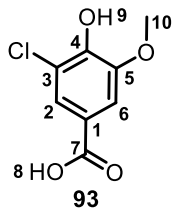


90

MP 122 – 126 °C. **Halogen analysis** provided by Raiford and Potter.[16] **IR (ATR, cm^{-1})** 3091, 3017, 2953 and 2846 (C-H str), 2230 (C≡N str), 1764 (C=O str), 1578, 1469 and 1407 (CH_3 bend), 1312, 1292, 1218, 1137, 1048 (C-O str), 859 (C-Cl str). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.36 (d, *J* = 1.7 Hz, 1H, $H_{3/5}$), 7.11 (d, *J* = 1.7 Hz, 1H, $H_{3/5}$), 3.88 (s, 3H, H_{10}), 2.38 (s, 3H, H_8). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 167.2 (C_7), 153.3, 141.0, 129.7, 125.8, 117.4, 114.0, 111.0, 56.8 (C_{10}), 20.4 (C_8). The molecular ion was not confirmed by HRMS, but the correct compound structure was inferred from the confirmed structures that follow.

4.5.3 Synthesis of 3-chloro-4-hydroxy-5-methoxybenzoic acid (93)

2-Chloro-4-cyano-6-methoxyphenyl acetate (**90**) (0.826 g, 3.66 mmol) was heated at reflux in a KOH solution (6 M, 10 mL) for 18 h. Thereafter, the cooled reaction mixture was placed on ice and was acidified with concentrated HCl to a pH of 3. The resultant precipitate was filtered off as the pure title compound (0.742 g, 3.66 mmol) (R_f = 0.46, 100% EtOAc) produced quantitatively as a white solid and was used without further purification).

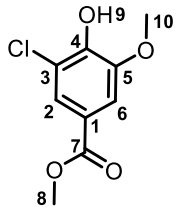


93

Decomposed at 250 °C. **Halogen analysis** provided by Raiford and Potter.[16] **IR (ATR, cm^{-1})** 3300 – 2350 (O-H str), 1674 (C=O str), 1596, 1421, 1299, 1237 (C-O str), 1108, 1046. **^{1}H NMR (400 MHz, DMSO-d) δ** 7.51 (d, *J* = 1.9 Hz, 1H, H_2), 7.41 (d, *J* = 1.9 Hz, 1H, H_6), 3.87 (s, 3H, H_{10}). **^{13}C NMR (101 MHz, DMSO-d) δ** 166.3 (C_7), 148.2, 147.1, 123.3, 121.8, 119.5, 111.0, 56.3 (C_{10}). **HRMS** calculated for $C_8H_6O_4Cl$ $[M-H]^-$, 200.9955, found 200.9964.

4.5.4 Synthesis of methyl 3-chloro-4-hydroxy-5-methoxybenzoate (94)

3-Chloro-4-hydroxy-5-methoxybenzoic acid (**93**) (0.500 g, 2.47 mmol), H_2SO_4 (0.2 eq., 0.03 mL, 0.6 mmol) and dry MeOH (15 mL) were heated together in a two-neck round-bottom flask, at reflux for 20 h. The reaction mixture was then added to ice-water (150 mL), neutralized with saturated $NaHCO_3$ solution and extracted with EtOAc (3 x 60 mL). The combined organic layers were dried over $MgSO_4$, filtered and the solvent was removed under reduced pressure. Purification was done by column chromatography (5 – 50% EtOAc/Hexane) to yield the title compound (0.301 g, 1.39 mmol, 88%) (R_f = 0.51, 50% EtOAc/Hexane) as a white solid.

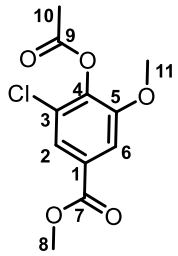


94

MP 130 – 132 °C. **IR (ATR, cm^{-1})** 3500 – 3150 (O-H str), 3015 and 2945 (C-H str), 1705 (C=O str), 1513, 1417, 1300 and 1225 (C-O str), 1185, 1107, 1059, 996, 857, 754 (C-Cl). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.72 (d, *J* = 1.8 Hz, 1H, H_2), 7.47 (d, *J* = 1.8 Hz, 1H, H_6), 6.23 (s, 1H, H_9), 3.96 (s, 3H, H_8), 3.90 (s, 3H, H_{10}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 166.1 (C_7), 147.1, 146.3, 124.5, 122.4, 119.6, 110.4, 56.7 (C_{10}), 52.4 (C_8). **HRMS** calculated for $C_9H_8O_4Cl$ $[M-H]^-$, 215.0111, found 215.0113.

4.5.5 Synthesis of methyl 4-acetoxy-3-chloro-5-methoxybenzoate (95)

Methyl 3-chloro-4-hydroxy-5-methoxybenzoate (**94**) (0.277 g, 1.28 mmol) and NaOAc (6.6 eq., 26.5 mg, 0.323 mmol) were heated at reflux in Ac_2O (5 mL) for 3 h. The reaction mixture was then poured into a beaker containing ice-water and was neutralized with KOH solution (1 M). Extraction was done with EtOAc (3 x 30 mL). The combined organic layers were dried over $MgSO_4$ and filtered, after which the solvent was removed under reduced pressure. Purification by column chromatography (5 – 50% EtOAc/Hexane) resulted in the title compound (0.278 g, 1.08 mmol, 92%) (R_f = 0.68, 50% EtOAc/Hexane) as a white solid.



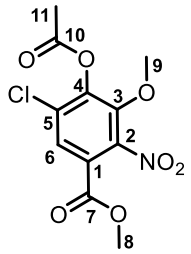
95

MP 74 – 76 °C. **Elemental analysis** provided in publication by Hodges and Taylor.[17] **IR (ATR, cm^{-1})** 3103 and 2954 (C-H str), 1771 and 1714 (C=O str), 1294 and 1247 (C-O str), 1163, 1041, 984, 764, 748 (C-Cl str). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.74 (d, *J* = 1.8 Hz, 1H, H_2), 7.54 (d, *J* = 1.8 Hz, 1H, H_6), 3.92 (s, 3H, H_8), 3.89 (s, 3H, H_{11}), 2.37 (s, 3H, H_{10}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 167.6 (C_9), 165.6 (C_7), 152.6, 140.6, 128.9, 128.4, 123.3, 111.8, 56.6 (C_{11}), 52.7 (C_8), 20.4 (C_{10}). **HRMS** calculated for $C_{11}H_{12}O_5Cl$ $[M+H]^+$, 259.0373, found 259.0374.

4.5.6 Synthesis of methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (96)

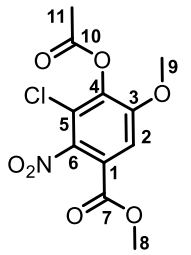
Fuming HNO_3 (5 mL) was cooled to -30 °C, after which methyl 4-acetoxy-3-chloro-5-methoxybenzoate (**95**) (272 mg, 1.05 mmol) was added in small portions, so as to keep the temperature between -30 and -20 °C. The reaction mixture was then stirred for 1 h. after addition,

after which it was added to ice-water and neutralized with a saturated $NaHCO_3$ solution. The organic compounds were subsequently extracted with EtOAc (3 x 30 mL), the combined organic layers dried over $MgSO_4$ and the organic products were purified by column chromatography (5 – 50% EtOAc/Hexane) to yield the title compound (247 mg, 0.814 mmol, 77%) (R_f = 0.63, 50% EtOAc/ Hexane) as a light-orange solid. Two by-products were obtained and were identified as methyl 4-acetoxy-3-chloro-5-methoxy-2-nitrobenzoate (**101**) (67.2 mg, 0.212 mmol, 21%) (R_f = 0.53, 50% EtOAc/Hexane) and methyl 4-acetoxy-3-chloro-5-methoxy-2,6-dinitrobenzoate (**102**) (7.34 mg, 0.0211 mmol, 2%) (R_f = 0.61, 50% EtOAC/Hexane). When the reaction mixture was left to stir for only 30 min. after the addition of compound **95**, a maximum yield of 86% of the title compound was obtained, while methyl 4-acetoxy-3-chloro-5-methoxy-2-nitrobenzoate (**101**) was obtained in 14% yield as the only by-product.



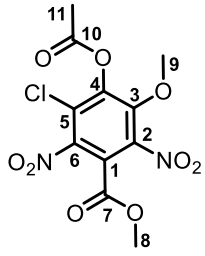
96

MP 87 – 89 °C. **Elemental analysis** provided in publication by Hodges and Taylor.[17] **IR (ATR, cm^{-1})** 3088, 3003 and 2957 (C-H str), 1779 and 1730 (C=O str), 1545 and 1310 (N=O str), 1153 and 1033 (C-O str), 900, 848, 759. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.88 (s, 1H, H_6), 3.92 (s, 3H, H_8), 3.91 (s, 3H, H_9), 2.43 (s, 3H, H_{11}). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 166.7 (C_7), 161.9, 146.1, 131.0, 126.9, 121.5, 63.4 (C_9), 53.6 (C_8), 20.5 (C_{11}). One carbon peak, presumably that of C_{10}, was not visible in the ^{13}C spectrum. **HRMS** calculated for $C_{11}H_9NO_7Cl$ $[M-H]^-$, instead $C_9H_7NO_6Cl$ was detected, 259.9962, found 259.9969.



101

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.50 (s, 1H, H_2), 3.95 (s, 3H, H_8), 3.92 (s, 3H, H_9), 2.39 (s, 3H, H_{11}).[17]



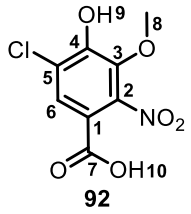
102

^{1}H NMR (300 MHz, $CDCl_3$) δ 4.01 (s, 3H, H_8), 3.90 (s, 3H, H_9), 2.47 (s, 3H, H_{11}).

It should be noted that the literature melting point of methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**) corresponds well to the measured meting point. The literature melting point of methyl 4-acetoxy-3-chloro-5-methoxy-2-nitrobenzoate (**101**) is significantly higher at 140 – 141 °C.[17]

4.5.7 Synthesis of 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (92)

Methyl 4-acetoxy-5-chloro-3-methoxy-2-nitrobenzoate (**96**) (0.126 g, 0.415 mmol) was heated at reflux in a MeOH (5 mL) and sodium hydroxide solution (4 M, 1 mL) for 18 h. The reaction mixture was then acidified with HCl solution (2 M) to a pH of 3. Extraction was done with EtOAc (3 x 30 mL), after which the combined organic layers were dried over $MgSO_4$, filtered and the solvent evaporated. The crude product, obtained as a light-yellow solid, was used as is in the next reaction.



92

Decomposed at 185 °C. **Elemental analysis** provided in publication by Hodges and Taylor.[17] **IR (ATR, cm^{-1})** 3576 and 3513 (C-H str), 3300 – 2200 (O-H str), 1690 (C=O str), 1542 and 1371 (N=O str), 1026 (C-O str), 815 (C-Cl str). **^{1}H NMR (300 MHz, DMSO-d) δ** 7.75 (s, 1H, H_6), 3.80 (s, 3H, H_8). H_9 and H_{10} were not visible in the ^{1}H NMR spectrum. **^{13}C NMR (75 MHz, DMSO-d) δ** 162.8 (C_7), 152.3, 140. 6, 127.0, 123.0, 113.2, 62.4 (C_8). **HRMS** calculated for $C_8H_5NO_6Cl$ $[M-H]^-$, 245.9805, found 245.9816.

4.5.8 Synthesis of 6-chloro-2-methoxy-3-nitrophenol (97)

A summary of the different decarboxylation methods attempted is given in the table below. The two successful methods (2 and 5) are described in detail below.

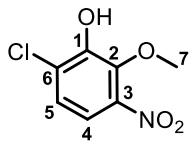
Method	Metal source	Ligand	Additive	Solvent	Temp. (°C)	Time (h.)	Result
1	Cu_2O	1,10-phenanthroline	-	NMP:Quinoline 3:1	170	12	8 spots on TLC plate; too little material collected for NMR spectroscopic analysis
2	AgOAc	-	K_2CO_3	NMP	120	16	3 spots on TLC plate; 14% yield over 2 steps
3	-	-	-	Diphenyl ether	258	<0.5	Product and solvent had the same R_f
4	Cu	-	-	Quinoline	238	4	No Product spot visible on TLC plate
5	Cu_2O	1,10-phenanthroline	-	NMP:Quinoline 3:1	170	1	48% yield over two steps

Method 2

To a two-neck round-bottom flask was added 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**) (0.103 g, 0.415 mmol), silver acetate (AgOAc) (0.1 eq., 6.92 mg, 0.0415 mmol), K_2CO_3 (0.15 eq., 8.60 mg, 0.0622 mmol) and NMP (4 mL). The reaction mixture was heated at 120 °C for 16 h., after which it was cooled to RT. H_2O (10 mL) was added and extraction was done with EtOAc (3 x 15 mL). The combined organic layers were dried over $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. After purification by column chromatography, the pure title product [11.6 mg, 0.0570 mmol, 14% (over two steps)] (R_f = 0.63, 50% EtOAc/Hexane) was obtained as a light-orange solid.

Method 5

To a Schlenk tube, 5-chloro-4-hydroxy-3-methoxy-2-nitrobenzoic acid (**92**) (248 mg, 1.00 mmol), copper(I) oxide (Cu_2O) (0.1 eq., 14.3 mg, 0.100 mmol) and 1,10-phenanthroline (0.1 eq., 18.0 mg, 0.100 mmol) was added and the Schlenk tube was flushed with alternating vacuum and nitrogen purge cycles. A solution of NMP (1.5 mL) and quinoline (0.5 mL) was then added using a syringe. The resulting mixture was stirred at 170 °C for 1 h., after which the solution was cooled to RT and poured into aqueous HCl (5 M, 2 mL). The aqueous layer was extracted with Et_2O (3 x 5 mL), after which the combined organic layers were further extracted with aqueous HCl (5 N, 5 x 5 mL) to remove quinoline, which co-elutes with the product during column chromatography. The organic layer was then washed with H_2O (5 mL) and brine (5 mL), dried over $MgSO_4$, filtered and the solvent was evaporated. Column chromatography (5 – 30% EtOAc/Hexane) was used as a means of purification and resulted in the pure title compound [98.6 mg, 0.484 mmol, 48% (over two steps)] (R_f = 0.63, 50% EtOAc/ Hexane) as a light-orange solid.

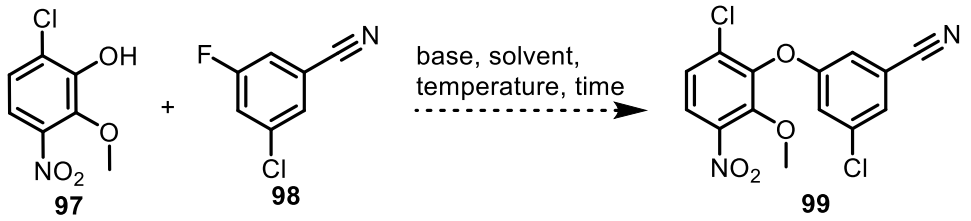


97

MP 86 – 88 °C. **Elemental analysis** provided in publication by Hodges and Taylor.[17] **IR (ATR, cm^{-1})** 3550 – 3200 (O-H str), 3103 and 2957 (C-H str), 1518 and 1335 (N=O str), 1200 (C-O str), 995, 812 (C-Cl). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.43 (d, *J* = 9.0 Hz, 1H, H_4), 7.21 (d, *J* = 9.0 Hz, 1H, H_5), 6.19 (s, 1H, OH), 4.02 (s, 3H, H_7). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 147.3, 142.3, 126.1, 124.4 (C_5), 116.6 (C_4), 62.7 (C_7). One carbon peak, presumably C_3 or C_6, was not visible in the ^{13}C NMR spectrum. **HRMS** calculated for $C_7H_5NO_4Cl$ [M-H]$^-$, 201.9907, found 201.9916.

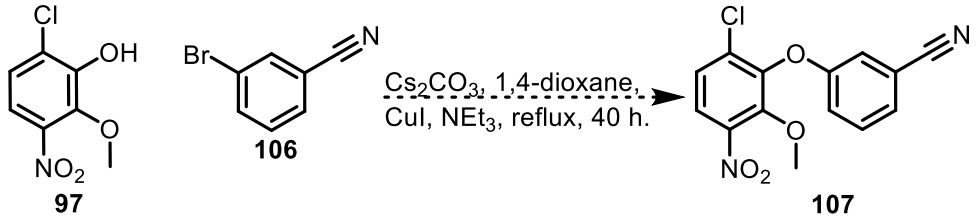
4.5.9 Attempted synthesis of 3-chloro-5-(6-chloro-2-methoxy-3-nitrophenoxy)benzonitrile (99)

A summary of the various conditions tried to achieve the S_NAr reaction, between 6-chloro-2-methoxy-3-nitrophenol (**97**) and 3-chloro-5-fluorobenzonitrile (**98**) (1.2 eq.), is given in the table below, none of which resulted in product. In all cases, the reaction was carried out in a dry, nitrogen-flushed Schlenk containing the indicated solvent to make up a 0.35 – 0.5 M solution.



#	Base	Solvent	Temperature (°C)	Time (h.)	Result
1	K_2CO_3 (2 eq.)	DMF	80	18	SM
2	K_2CO_3 (2 eq.)	DMF	100	18	SM
3	Cs_2CO_3 (1.2 eq.)	MeCN	rt	2	SM
4	Cs_2CO_3 (1.2 eq.)	MeCN	50	20	SM
5	Cs_2CO_3 (1.2 eq.)	MeCN	100	20	SM
6	K_2CO_3 (2 eq.)	NMP	140	40	3 unidentified compounds

4.5.10 Attempted synthesis of 3-(6-chloro-2-methoxy-3-nitrophenoxy) benzonitrile (107)

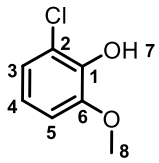


To a solution of 6-chloro-2-methoxy-3-nitrophenol (**97**) (36.3 mg, 0.178 mmol) and 3-bromobenzonitrile (**106**) (1 eq., 32.5 mg, 0.178 mmol) in 1,4-dioxane (1.5 mL), CuI (0.1 eq., 3.40 mg, 0.0178 mmol), NEt_3 (0.1 eq., 1.80 mg, 0.0178 mmol) and Cs_2CO_3 (2 eq., 0.116 g, 0.357 mmol) were added. The reaction mixture was then heated at reflux for 40 h. When no reaction was observed on the TLC plate, the solvent was removed under reduced pressure, the residue dissolved in EtOAc (10 mL) and was then washed sequentially with HCl (1 M, 5 x 5 mL). The organic layer was subsequently washed with saturated $NaHCO_3$ solution (5 mL) and then with brine (5 mL), after which it was dried over $MgSO_4$, filtered and concentrated under reduced pressure. The starting materials [6-chloro-2-methoxy-3-nitrophenol (**97**) (34.6 mg, 0.170 mmol) and 3-bromobenzonitrile (**106**) (27.4 mg, 0.176 mmol)] were recovered after purification with column chromatography.

4.5.11 Synthesis of 2-chloro-6-methoxyphenol (110)

3-Chloro-4-hydroxy-5-methoxybenzoic acid (**93**) (0.102 mg, 0.505 mmol), Cu_2O (0.1 eq., 7.23 mg, 0.0505 mmol) and 1,10-phenanthroline (0.1 eq., 9.10 mg, 0.0505 mmol) were added to an oven dried Schlenk tube, after which it was flushed with alternating vacuum and nitrogen purge cycles.

NMP (1 mL) and quinoline (0.3 mL) were then added using a syringe and the resulting mixture was stirred at 170 °C for 1 h. The solution was then poured into aqueous HCl (5 N, 2 mL) and the aqueous layer was extracted with Et_2O (3 x 5 mL), after which the combined organic layers were further extracted with aqueous HCl (5 N, 5 x 5 mL), to remove quinoline which co-elutes with the product when columned. The organic layer was then washed with H_2O (5 mL) and brine (5 mL), dried over $MgSO_4$, filtered and the solvent was evaporated. The crude material was purified by means of column chromatography (5 – 30% EtOAc/Hexane) and the title compound (57.3 mg, 0.361 mmol, 72%) (R_f = 0.64, 50% EtOAc/Hexane) was obtained as a dark orange solid.



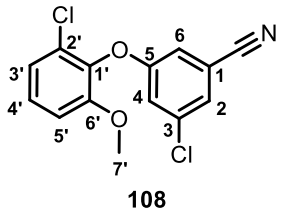
110

^{1}H NMR (300 MHz, $CDCl_3$) δ 6.95 (dd, *J* = 6.3, 3.3 Hz, 1H, H_5), 6.85 – 6.73 (m, 2H, H_3 and H_4), 5.83 (s, 1H, H_7), 3.90 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 147.7, 142.2, 122.2, 120.1, 119.7, 109.4, 56.5 (C_8).

The spectra for this compound matched those reported in the literature.[36]

4.5.12 Synthesis of 3-chloro-5-(2-chloro-6-methoxyphenoxy)benzonitrile (108)

2-Chloro-6-methoxyphenol (**110**) (37.8 mg, 0.238 mmol), 3-chloro-5-fluorobenzonitrile (**98**) (1.2 eq., 44.5 mg, 0.286 mmol) and K_2CO_3 (2 eq., 66.0 mg, 0.478 mmol) was added to an oven dried Schlenk tube, which was then subjected to alternating vacuum and nitrogen purge cycles. DMF (1 mL) was added with a syringe and the reaction was stirred at 80 °C for 67 h. H_2O (5 mL) was then added to the cooled reaction mixture and the aqueous layer was extracted with EtOAc (3 x 5 mL). The crude material was purified by column chromatography (5 – 30% EtOAc/Hexane) to yield the title compound (12.0 mg, 0.0408 mmol, 17%) (R_f = 0.48, 20% EtOAc/Hexane) as an off-white solid.



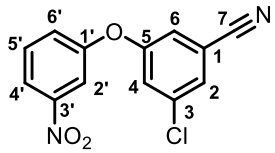
108

MP 116 – 118 °C. **IR (ATR, cm^{-1})** 3067, 2918 and 2849 (C-H str), 2235 (C≡N str), 1590, 1572, 1480, 1260 and 1042 (C-O str), 857 and 770 (C-Cl). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.30 (dd [app. t], J_1 = J_2 = 1.4 Hz, 1H, H_6), 7.22 (dd [app. t], J_1 = J_2 = 8.3 Hz, 1H, $H_{4'}$), 7.12 – 7.09 (m, 2H, 2 x ArH), 7.01 – 6.88 (m, 2H, 2 x ArH), 3.80 (s, 3H, $H_{7'}$). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 158.5, 136.3, 127.3, 125.8, 122.5, 120.8, 117.0, 114.4, 111.4, 56.4 ($C_{7'}$). Note that due to the small quantity isolated, ^{13}C NMR spectroscopic analysis proved to be a difficult task. As a result, only 10 of the 14 expected signals were detected. **HRMS** calculated for $C_{14}H_{10}NO_2Cl_2$ $[M+H]^+$, 294.0089, found 294.0079.

4.5.13 Synthesis of 3-chloro-5-(3-nitrophenoxy)benzonitrile (109)

3-Nitrophenol (**111**) (50.0 mg, 0.359 mmol), 3-chloro-5-fluorobenzonitrile (**98**) (1.2 eq., 67.1 mg, 0.431 mmol) and K_2CO_3 (2 eq., 99.3 mg, 0.719 mmol) were added to a Schlenk tube, which was then flushed with alternating vacuum and nitrogen purge cycles. DMF (2 mL) was added and the reaction mixture was stirred at 80 °C for 65 h. After cooling the mixture to RT, H_2O (5 mL) and EtOAc

(5 mL) were added and the aqueous layer was separated. The aqueous layer was further extracted with EtOAc (3 x 5 mL) and the combined organic layers were dried over $MgSO_4$, filtered and evaporated. Purification by column chromatography (5 – 50% EtOAc/Hexane) yielded the title compound (67.7 mg, 0.246 mmol, 67%) (R_f = 0.45, 20% EtOAc/Hexane) as a light-yellow solid.

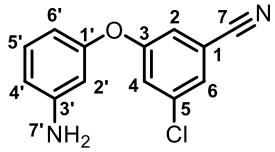


109

MP 100 – 102 °C. **IR (ATR, cm^{-1})** 3084 (C-H str), 2237 (C≡N str), 1579, 1523 and 1347 (N=O str), 1265 (C-O str), 1211, 977, 860. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.10 (ddd, *J* = 8.2, 2.3, 1.0 Hz, 1H, $H_{4'}$), 7.88 (dd [app. t], $J_1 = J_2$ = 2.2 Hz, 1H, H_6), 7.61 (dd [app. t], $J_1 = J_2$ = 8.2 Hz, 1H, $H_{5'}$), 7.45 (dd [app. t], $J_1 = J_2$ = 2.2, 1H, H_2), 7.39 (ddd, *J* = 8.2, 2.3, 1.0 Hz, 1H, $H_{6'}$), 7.26 (dd [app. t], $J_1 = J_2$ = 2.2 Hz, 1H, H_4), 7.18 (dd, *J* = 2.3, 1.0 Hz, 1H, $H_{2'}$). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 157.6, 156.1, 137.0, 131.3, 127.7, 125.6, 123.8, 120.3, 120.0, 116.8, 115.3, 114.8. Note that the molecular ion was not confirmed with HRMS, but that the correct structure was inferred from the product that follows.

4.5.14 Synthesis of 3-(3-aminophenoxy)-5-chlorobenzonitrile (112)

3-Chloro-5-(3-nitrophenoxy)benzonitrile (**109**) (63.7 mg, 0.232 mmol), followed by Fe powder (5.2 eq., 67.3 mg, 1.21 mmol) were added to a mixture of glacial acetic acid (2 mL), EtOH (2 mL) and H_2O (1 mL). The suspension was exposed to ultrasonic irradiation for 3 h. and 20 min., after which the mixture was filtered and washed with EtOAc (3 x 10 mL). The filtrate was partitioned with a KOH solution (2 M, 30 mL) and the aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (2 x 20 mL) and H_2O (3 x 30 mL) and dried over $MgSO_4$. After filtration and concentration under reduced pressure, the crude material was subjected to column chromatography (5 – 30% EtOAc/Hexane) from which the title compound **112** was obtained (18.3 mg, 0.0748 mmol, 32%) (R_f = 0.26, 20% EtOAc/Hexane) as a yellow oil.



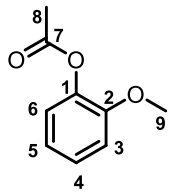
112

IR (ATR, cm^{-1}) 3465, 3375 (N-H str), 3078 (C-H str), 2234 (C≡N str), 1567, 1489, 1424, 1267 (C-O str), 996, 983, 855, 777. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.30 (dd [app. t], $J_1 = J_2$ = 2.0 Hz, 1H, H_2), 7.19 (dd [app. t], $J_1 = J_2$ = 2.1 Hz, 1 H, $H_{2'}$), 7.16 (dd [app. t], $J_1 = J_2$ = 8.0 Hz, 1H, $H_{5'}$), 7.11 (dd, *J* = 2.0, 1.3 Hz, 1H, H_4), 6.54 (ddd, *J* = 8.0, 2.1, 0.6 Hz, 1H, $H_{4'}$), 6.39 (ddd, *J* = 8.0, 2.1, 0.6 Hz, 1H, $H_{6'}$), 6.34 (dd [app. t], $J_1 = J_2$ = 2.0 Hz, 1H, H_6), 3.80 (s, 2H, $H_{7'}$). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 159.3, 155.9, 148.7, 136.3, 131.1, 125.8, 122.8, 119.4, 117.3, 114.5, 112.2, 109.8, 106.7. **HRMS** calculated for $C_{13}H_{10}N_2OCl$ $[M+H]^+$, 245.0482, found 245.0484.

4.5.15 Synthesis of 2-methoxyphenyl acetate (114)

AcCl (1.16 eq., 1.50 mL, 21.0 mmol) was added dropwise to stirred solution of commercially available guaiacol (**113**) (2.00 mL, 17.9 mmol) and pyridine (2.2 eq., 3.20 mL, 39.7 mmol) in DCM (15 mL), maintained at 0 °C under N_2 atmosphere. The reaction mixture was allowed to warm to RT over the period of 1 h. and stirring was continued for 22 h. The solution was then poured into ice-

cooled H_3PO_4 (1 M, 40 mL) and the separated aqueous phase was extracted with DCM (3 x 20 mL). The combined organic extracts were washed with brine (1 x 50 mL), dried on $MgSO_4$, filtered and concentrated under reduced pressure. Column chromatography (5 – 30% EtOAc/Hexane) yielded the pure title compound **114** (2.80 g, 16.8 mmol, 94%) (R_f = 0.46, 20% EtOAc/Hexane) as a clear oil.



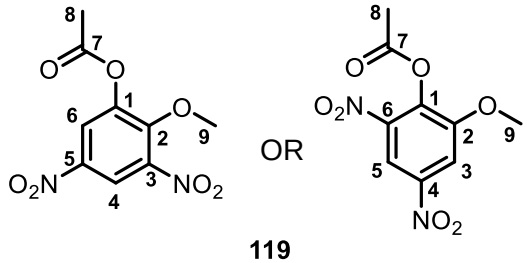
114

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.21 (ddd, *J* = 8.3, 7.3, 1.8 Hz, 1H, H_4), 7.04 (dd, *J* = 7.8, 1.8 Hz, 1H, H_6), 7.01 – 6.91 (m, 2H, H_3 and H_5), 3.83 (s, 3H, H_9), 2.32 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 169.2 (C_7), 151.2, 139.9, 127.0, 122.9, 120.9, 112.5, 56.0 (C_9), 20.8 (C_8).

The spectra for this compound matched those reported in the literature.[48]

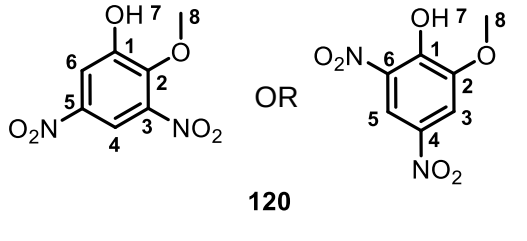
4.5.16 Attempted synthesis of 2-methoxy-3-nitrophenyl acetate (115)

Fuming HNO_3 (5 mL) was cooled to -30 °C before 2-methoxyphenyl acetate (**114**) (0.30 mL, 2.0 mmol) was added portion-wise so as to keep the temperature between -30 and -20 °C. After complete addition of compound **114**, the reaction mixture was stirred for a further 30 min. The reaction mixture was then poured onto ice and a saturated $NaHCO_3$ solution was used to neutralise the solution. Extraction was done with EtOAc (3 x 30 mL) and the combined organic layers were dried over $MgSO_4$. Two products, mostly isolated as a mixture, were obtained. The first product, with a R_f of 0.63 (50% EtOAc/Hexane), was identified as a dinitrophenyl acetate, while the second product was identified as a dinitrophenol, with an R_f of 0.53 (50% EtOAc/Hexane).



119

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.60 (d, *J* = 2.8 Hz, 1H, $H_{4/5}$), 8.23 (d, *J* = 2.8 Hz, 1H, $H_{6/3}$), 4.05 (s, 3H, H_9), 2.43 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 167.8 (C_7), 151.4, 145.3, 123.4, 118.5, 63.1 (C_9), 20.7 (C_8). Note that two carbon peaks, presumably the two C-NO_2 carbons peaks, were not visible in the ^{13}C NMR spectrum.



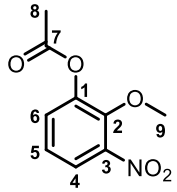
120

^{1}H NMR (300 MHz, $CDCl_3$) δ 8.35 (d, *J* = 2.7 Hz, 1H, $H_{4/5}$), 8.06 (d, *J* = 2.7 Hz, 1H, $H_{6/3}$), 6.41 (s, 1H, H_7), 4.07 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 151.2, 146.2, 114.9, 112.9, 63.2 (C_8). Note that two carbon peaks, presumably the two C-NO_2 carbons peaks, were not visible in the ^{13}C NMR spectrum.

4.5.17 Synthesis of 2-methoxy-3-nitrophenyl acetate (115)

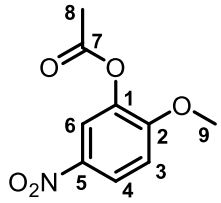
A stirred solution of guaiacol acetate (**114**) (0.10 mL, 0.68 mmol) in Ac_2O (6.2 eq., 0.40 mL, 4.2 mmol) at 0 °C was treated dropwise with an ice-cooled solution of fuming HNO_3 (7 eq., 0.20 mL, 4.8 mmol)

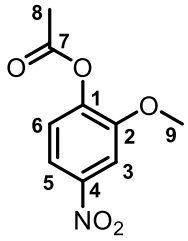
in Ac_2O (0.2 eq., 0.30 mL, 3.2 mmol), prepared by mixing the reagents at RT. The mixture was then added to a saturated $NaHCO_3$ solution and ice. $NaHCO_3$ was added until a neutral solution was obtained. Extraction was done with EtOAc (3 x 30 mL) and the combined organic layers were dried over $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. The pure title compound **115** (0.105 g, 0.495 mmol, 73%) (R_f = 0.32, 20% EtOAc/Hexane) was obtained after column chromatography (5 – 30% EtOAc/Hexane) as a yellow oil. The by-product, either 2-methoxy-5-nitrophenyl acetate or 2-methoxy-4-nitrophenyl acetate (**121**) (0.0319 g, 0.151 mmol, 22%) (R_f = 0.21, 20% EtOAc/Hexane), was obtained as a yellow solid.



115

IR (ATR, cm^{-1}) 3125, 3094, 2923 and 2851 (C-H str), 1760 (C=O str), 1595 and 1340 (N=O str), 1503, 1295 and 1180 (C-O str), 1140, 1082, 1015, 904, 819, 804, 742 (C-Cl). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.72 (dd, *J* = 8.1, 1.7 Hz, 1H, H_4), 7.34 (dd, *J* = 8.1, 1.7 Hz, 1H, H_6), 7.21 (dd [app. t], $J_1 = J_2$ = 8.1 Hz, 1H, H_5), 3.94 (s, 3H, H_9), 2.38 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 168.5 (C_7), 146.4, 145.5, 128.4, 123.7, 122.7, 62.7 (C_9), 20.8 (C_8). Note that one carbon peak, presumably the C-NO_2 peak, was not visible in the ^{13}C NMR spectrum. **HRMS** calculated for $C_9H_{10}NO_5$ $[M+H]^+$ 212.0559, found 212.0562.



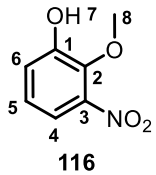
OR


121

MP 90 – 92 °C. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.17 (dd, *J* = 9.1, 2.7 Hz, 1H, $H_{4/5}$), 7.97 (d, *J* = 2.7 Hz, 1H, $H_{6/3}$), 7.03 (d, *J* = 9.1 Hz, 1H, $H_{3/6}$), 3.94 (s, 3H, H_9), 2.34 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 168.4 (C_7), 156.8, 141.2, 139.5, 123.4, 119.3, 111.5, 56.7 (C_9), 20.6 (C_8).

4.5.18 Synthesis of 2-methoxy-3-nitrophenol (116)

To 3-nitroguaiacol acetate (**115**) (93.0 mg, 0.440 mmol), methanol (4.5 mL) and a NaOH solution (4 N, 0.5 mL) was added. The reaction mixture was heated at reflux for 3 h., after which a HCl solution (2 M, 5 mL) was added and the aqueous layer was extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated under reduced pressure. After column chromatography (5 – 40% EtOAc/Hexane), the pure title compound **116** (67.8 mg, 0.401 mmol, 91%) (R_f = 0.59, 50% EtOAc/Hexane) was obtained as a light-yellow solid.



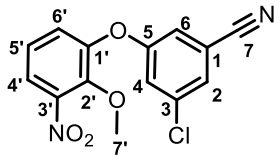
116

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.47 (dd, *J* = 8.2, 1.7 Hz, 1H, $H_{4/6}$), 7.23 (dd, *J* = 8.2, 1.7 Hz, 1H, $H_{4/6}$), 7.11 (dd [app. t], $J_1 = J_2$ = 8.2 Hz, 1H, H_5), 5.99 (s, 1H, H_7), 3.97 (s, 3H, H_8). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 150.7, 141.3, 124.5, 120.8, 117.1, 62.7 (C_8). Note that one carbon peak, presumably the C-NO_2 peak, was not visible in the ^{13}C NMR spectrum.

The spectra for this compound matched those reported in the literature.[49]

4.5.19 Synthesis of 3-chloro-5-(2-methoxy-3-nitrophenoxy)benzonitrile (117)

A Schlenk tube containing 2-methoxy-3-nitrophenol (**116**) (0.100 g, 0.591 mmol), 3-chloro-5-fluorobenzonitrile (**98**) (1.2 eq., 0.110 g, 0.710 mmol) and K_2CO_3 (2 eq., 0.163 g, 1.18 mmol) was flushed with alternating vacuum and nitrogen purge cycles, after which DMF (4 mL) was added with a syringe. The reaction mixture was heated at 80 °C for 48 h. H_2O (10 mL) was then added and the aqueous layer was extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtered and evaporated under reduced pressure. Column chromatography (1 – 5% EtOAc/Hexane) was used as a purification method and the pure title compound **117** (42.3 mg, 0.139 mmol, 23%) (R_f = 0.44, 20% EtOAc/Hexane) was obtained as an off-white semi-solid.

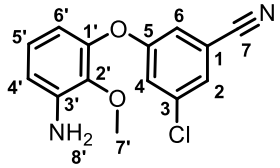


117

IR (ATR, cm^{-1}) 3061, 2918 and 2849 (C-H str), 2236 (C≡N str), 1528 and 1270 (N=O str), 1424, 1347, 983, 860, 745 (C-Cl). **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.73 (dd, *J* = 7.6, 2.2 Hz, 1H, ArH), 7.38 (dd [app.t], J_1 = J_2 = 1.3 Hz, 1H, ArH), 7.34 – 7.29 (m, 1H, ArH), 7.28 – 7.23 (m, 1H, ArH), 7.15 (dd [app.t], J_1 = J_2 = 2.2 Hz, 1H, ArH), 7.10 – 6.06 (m, 1H, ArH), 3.95 (s, 3H, $H_{7'}$). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 158.1, 148.7, 146.2, 136.8, 127.0, 126.7, 124.6, 122.2, 121.8, 118.4, 116.9, 115.0, 62.7 ($C_{7'}$). Note that the molecular ion was not confirmed with HRMS, but that the correct structure was inferred from the subsequent amine product.

4.5.20 Synthesis of 3-(3-amino-2-methoxyphenoxy)-5-chlorobenzonitrile (118)

To 3-chloro-5-(2-methoxy-3-nitrophenoxy)benzonitrile (**117**) (15.4 mg, 0.0505 mmol) and Fe powder (14.7 mg, 0.263 mmol) was added a mixture of glacial acetic acid (2 mL), EtOH (2 mL) and H_2O (1 mL). The suspension was exposed to ultrasonic irradiation for 3 h. and 40 min. The reaction mixture was then filtered and washed with EtOAc (3 x 10 mL). The filtrate was partitioned with a KOH solution (2 M, 30 mL) and the aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (2 x 20 mL) and H_2O (3 x 30 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The crude material was subjected to column chromatography (5 – 30% EtOAc/Hexane) and the pure title compound **118** was obtained (7.60 mg, 0.0277 mmol, 55%) (R_f = 0.24, 20% EtOAc/Hexane) as a yellow solid.



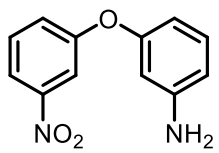
118

MP 109 – 110 °C. **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.29 (dd [app. t], J_1 = J_2 = 1.3 Hz, 1H, H_6), 7.16 (dd [app. t], J_1 = J_2 = 2.2 Hz, 1H, H_4), 7.07 (dd, *J* = 2.2, 1.3 Hz, 1H, H_2), 6.92 (dd [app. t], J_1 = J_2 = 8.1 Hz, 1H, $H_{5'}$), 6.65 (dd, *J* = 8.1, 1.3 Hz, 1H, $H_{4'}$), 6.41 (dd, *J* = 8.1, 1.3 Hz, 1H, $H_{6'}$), 3.97 (s, 2H, $H_{8'}$), 3.76 (s, 3H, $H_{7'}$). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 159.3, 147.1, 142.0, 139.0, 136.3, 125.6, 125.1, 121.5, 118.1, 117.4, 114.5, 113.3, 111.4, 60.5 ($C_{7'}$). **HRMS** calculated for $C_{14}H_{12}N_2O_2Cl$

$[M+H]^+$, 275.0587, found 275.0582. Note that IR data was not obtained due to insufficient material left after biological testing.

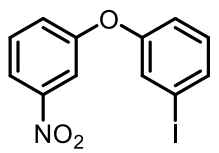
4.5.21 Additional attempted syntheses of diaryl ether compounds through S_NAr reactions

4.5.21.1 Attempted synthesis of 3-(3-nitrophenoxy)aniline (123)



3-Nitrophenol (**111**) (50.0 mg, 0.359 mmol), 3-fluoroaniline (**122**) (1.2 eq., 0.04 mL, 0.4 mmol) and K_2O_3 (2 eq., 99.4 mg, 0.719 mmol) were added to a Schlenk tube, which was subsequently flushed with alternating vacuum and nitrogen purge cycles. DMF (2 mL) was then added and the reaction mixture was stirred at 80 °C for 42 h. When no reaction was detected upon TLC analysis, the reaction mixture was cooled to RT, H_2O was added and extraction was done with EtOAC (3 x 10 mL). The combined organic layers were then dried over $MgSO_4$, filtered and concentrated under reduced pressure. Purification by column chromatography (5 – 50% EtOAc/Hexane) resulted in the recovery of the two starting materials, 3-nitrophenol (**111**) (49.5 mg, 0.356 mmol) and 3-fluoroaniline (**122**) (47.2 mg, 0.425 mmol).

4.5.21.2 Attempted synthesis of 1-iodo-3-(3-nitrophenoxy)benzene (125)

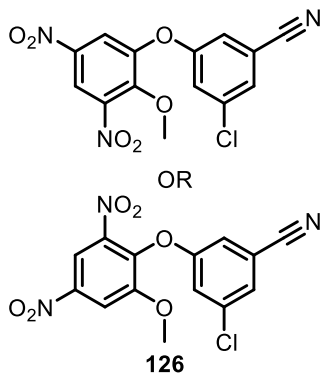


The same procedure was used as per the attempted synthesis of 3-(3-nitrophenoxy)aniline (**123**).

The equivalents used were as follows: 3-nitrophenol (**111**) (50.0 mg, 0.359 mmol), 3-fluoroiodobenzene (**124**) (1.2 eq., 0.05 mL, 0.4 mmol), K_2CO_3 (2 eq., 99.4 mg, 0.719 mmol), DMF (2 mL).

The reaction was not successful and the two starting materials, 3-nitrophenol (**111**) (49.2 mg, 0.354 mmol) and 3-fluoroiodobenzene (**124**) (86.6 mg, 0.390 mmol), were recovered after column chromatography (5 – 30% EtOAc/Hexane).

4.5.21.3 Attempted synthesis of 3-chloro-5-[2-methoxy-3,5 (or 4,6)-dinitrophenoxy]benzonitrile (126)



The deacetylated product (**120**) from the attempted synthesis of 2-methoxy-3-nitrophenyl acetate (**115**) (Section 4.5.16) (0.102 g, 0.475 mmol) was added to a Schlenk tube with 3-chloro-5-fluorobenzonitrile (**98**) (1.2 eq., 88.7 mg, 0.571 mmol) and K_2CO_3 (2 eq., 0.131 g, 0.951 mmol). The Schlenk tube was then flushed with alternating vacuum and nitrogen purge cycles, before DMF (3 mL) was added with a syringe. When after 24 h. at 80 °C, no reaction was observed on the TLC plate, the temperature was increased to 100 °C for a further 24 h.

The reaction mixture was then cooled to RT and H_2O (10 mL) was added. The organic products were extracted with EtOAc (3 x 10 mL) and the combined organic layers were dried over $MgSO_4$, filtered and concentrated. The two starting materials [**120** (0.101 g, 0.470 mmol) and 3-chloro-5-fluorobenzonitrile (**98**) (85.9 mg, 0.552 mmol)] were recovered after column chromatography (5 – 30% EtOAc/Hexane).

4.6 Bibliography

1 Z. K. Sweeney, J. P. Dunn, Y. Li, G. Heilek, P. Dunten, T. R. Elworthy, X. Han, S. F. Harris, D. R. Hirschfeld, J. H. Hogg, W. Huber, A. C. Kaiser, D. J. Kertesz, W. Kim, T. Mirzadegan, M. G. Roepel, Y. D. Saito, T. M. P. C. Silva, S. Swallow, J. L. Tracy, A. Villasenor, H. Vora, A. S. Zhou and K. Klumpp, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4352–4354.

2 Z. K. Sweeney, S. Acharya, A. Briggs, J. P. Dunn, T. R. Elworthy, J. Fretland, A. M. Giannetti, G. Heilek, Y. Li, A. C. Kaiser, M. Martin, Y. D. Saito, M. Smith, J. M. Suh, S. Swallow, J. Wu, J. Q. Hang, A. S. Zhou and K. Klumpp, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4348–4351.

3 Z. K. Sweeney, J. J. Kennedy-Smith, J. Wu, N. Arora, J. R. Billedeau, J. P. Davidson, J. Fretland, J. Q. Hang, G. M. Heilek, S. F. Harris, D. Hirschfeld, P. Inbar, H. Javanbakht, J. A. Jernelius, Q. Jin, Y. Li, W. Liang, R. Roetz, K. Sarma, M. Smith, D. Stefanidis, G. Su, J. M. Suh, A. G. Villaseñor, M. Welch, F. Zhang and K. Klumpp, *ChemMedChem*, 2009, **4**, 88–99.

4 Z. K. Sweeney, S. F. Harris, N. Arora, H. Javanbakht, Y. Li, J. Fretland, J. P. Davidson, J. R. Billedeau, S. K. Gleason, D. Hirschfeld, J. J. Kennedy-Smith, T. Mirzadegan, R. Roetz, M. Smith, S. Sperry, J. M. Suh, J. Wu, S. Tsing, A. G. Villasenor, A. Paul, G. Su, G. Heilek, J. Q. Hang, A. S. Zhou, J. A. Jernelius, F.-J. Zang and K. Klumpp, *J. Med. Chem.*, 2008, **51**, 7449–7458.

5 P. Vernazza, C. Wang, A. Pozniak, E. Weil, P. Pulik, D. A. Cooper, R. Kaplan, A. Lazzarin, H. Valdez, J. Goodrich, J. Mori, C. Craig and M. Tawadrous, *J. Acquir. Immune Defic. Syndr.*, 2013, **62**, 171–179.

6 R. Corbau, J. Mori, C. Phillips, L. Fishburn, A. Martin, C. Mowbray, W. Panton, C. Smith-Burchnell, A. Thornberry, H. Ringrose, T. Knöchel, S. Irving, M. Westby, A. Wood and M. Perros, *Antimicrob. Agents Chemother.*, 2010, **54**, 4451–4463.

7 C. E. Mowbray, C. Burt, R. Corbau, M. Perros, I. Tran, P. A. Stupple, R. Webster and A. Wood, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 5599–5602.

8 T. Fujiwara, A. Sato, M. El-Farrash, S. Miki, K. Abe, Y. Isaka, M. Kodama, Y. Wu, L. B. Chen, H. Harada, H. Sugimoto, M. Hatanaka and Y. Hinuma, *Antimicrob. Agents Chemother.*, 1998, **42**, 1340–1345.

9 T. Ohkawa, S. Goto, S. Miki, A. Sato, T. Kuroda, K. Iwatani, M. Takeuchi and M. Nakano, *Xenobiotica*, 1998, **28**, 877–886.

10 T. J. Tucker, J. T. Sisko, R. M. Tynebor, T. M. Williams, P. J. Felock, J. A. Flynn, M. Lai, Y.

Liang, G. McGaughey, M. Liu, G. Moyer, V. Munshi, R. Perlow-Poehnelt, S. Prasad, J. C. Reid, R. Sanchez, M. Torrent, J. P. Vacca, B.-L. Wan and Y. Yan, *J. Med. Chem.*, 2008, **51**, 6503–6511.

11 B. Côté, J. D. Burch, E. Asante-Appiah, C. Bayly, L. Bédard, M. Blouin, L. C. Campeau, E. Cauchon, M. Chan, A. Chefson, N. Coulombe, W. Cromlish, S. Debnath, D. Deschênes, K. Dupont-Gaudet, J. P. Falgueyret, R. Forget, S. Gagné, D. Gauvreau, M. Girardin, S. Guiral, E. Langlois, C. S. Li, N. Nguyen, R. Papp, S. Plamondon, A. Roy, S. Roy, R. Seliniotakis, M. St-Onge, S. Ouellet, P. Tawa, J. F. Truchon, J. Vacca, M. Wrona, Y. Yan and Y. Ducharme, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 917–922.

12 P. Zhan, X. Liu, Z. Li, C. Pannecouque and E. De Clercq, *Curr. Med. Chem.*, 2009, **16**, 3903–3917.

13 T. J. Tucker, S. Saggar, J. T. Sisko, R. M. Tynebor, T. M. Williams, P. J. Felock, J. A. Flynn, M.-T. Lai, Y. Liang, G. McGaughey, M. Liu, M. Miller, G. Moyer, V. Munshi, R. Perlow-Poehnelt, S. Prasad, R. Sanchez, M. Torrent, J. P. Vacca, B. L. Wan and Y. Yan, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 2959–2966.

14 J. J. Kennedy-Smith, N. Arora, J. R. Billedeau, J. Fretland, J. Q. Hang, G. M. Heilek, S. F. Harris, D. Hirschfeld, H. Javanbakht, Y. Li, W. Liang, R. Roetz, M. Smith, G. Su, J. M. Suh, A. G. Villaseñor, J. Wu, D. Yasuda, K. Klumpp and Z. K. Sweeney, *Med. Chem. Commun.*, 2010, **1**, 79.

15 S. J. Smerdon, J. Jäger, J. Wang, L. A. Kohlstaedt, A. J. Chirino, J. M. Friedman, P. A. Rice and T. A. Steitz, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 3911–3915.

16 L. C. Raiford and D. J. Potter, *J. Am. Chem. Soc.*, 1933, **55**, 1682–1685.

17 R. Hodges and A. Taylor, *J. Am. Chem. Soc.*, 1953, **0**, 4310–4314.

18 D. T. Mowry, *Chem. Rev.*, 1948, **42**, 189–283.

19 J. Pelouze, *Ann. der Pharm.*, 1834, **10**, 249.

20 H. Fehling, *Ann. der Chemie und Pharm.*, 1844, **49**, 91–97.

21 H. A. Curtis, *J. Chem. Educ.*, 1942, **19**, 161.

22 S. Gabriel and R. Meyer, *Chem. Ber.*, 1881, **14**, 1332–2341.

23 M. Lehr, *Arch. Pharm.*, 1996, **329**, 386–392.

24 V. Y. Kukushkin and A. J. L. Pombeiro, *Inorganica Chim. Acta*, 2005, **358**, 1–21.

25 B. S. Furniss, A. J. Hannaford, P. W. G. Smith and A. R. Tatchell, in *Vogel's Textbook of Practical Organic Chemistry*, Longman Scientific & Technical, London, Fith, 1989, pp. 695–697.

26 E. Fischer and A. Speier, *Chem. Ber.*, 1895, **28**, 3252–3258.

27 J. Clayden, N. Greevs, S. Warren and P. Wothers, *Organic Chemistry*, Oxford University Press, Oxford, Second, 2001.

28 J. G. Hoggett, R. B. Moodie, J. R. Penton and K. Schofield, *Nitration and Aromatic Reactivity.*,

Cambridge University Press, London, 1971.

29 L. J. Gooßen, W. R. Thiel, N. Rodríguez, C. Linder and B. Melzer, *Adv. Synth. Catal.*, 2007, **349**, 2241–2246.

30 L. J. Gooßen, C. Linder, N. Rodríguez, P. P. Lange and A. Fromm, *Chem. Commun.*, 2009, 7173.

31 J. L. Henderson, A. Sawant-Basak, J. B. Tuttle, A. B. Dounay, L. A. McAllister, J. Pandit, S. Rong, X. Hou, B. M. Bechle, J.-Y. Kim, V. Parikh, S. Ghosh, E. Evrard, L. E. Zawadzke, M. A. Salafia, B. Rago, R. S. Obach, A. Clark, K. R. Fonseca, C. Chang and P. R. Verhoest, *Med. Chem. Commun.*, 2013, **4**, 125–129.

32 R. Gomez, S. J. Jolly, T. Williams, J. P. Vacca, M. Torrent, G. McGaughey, M.-T. Lai, P. Felock, V. Munshi, D. DiStefano, J. Flynn, M. Miller, Y. Yan, J. Reid, R. Sanchez, Y. Liang, B. Paton, B.-L. Wan and N. Anthony, *J. Med. Chem.*, 2011, **54**, 7920–7933.

33 F. Ullmann and P. Sponagel, *Berichte der Dtsch. Chem. Gesellschaft*, 1905, **38**, 2211–2212.

34 F. Monnier and M. Taillefer, *Angew. Chem., Int. Ed.*, 2009, **48**, 6954–6971.

35 J. A. Olson and K. M. Shea, *Acc. Chem. Res.*, 2011, **44**, 311–321.

36 E. Vitaku and J. T. Njardarson, *Eur. J. Org. Chem.*, 2016, **2016**, 3679–3683.

37 L. Wang, P. Li, Z. Wu, J. Yan, M. Wang and Y. Ding, *Synthesis*, 2003, **13**, 2001–2004.

38 G. A. Heropoulos, S. Georgakopoulos and B. R. Steele, *Tetrahedron Lett.*, 2005, **46**, 2469–2473.

39 K. Nomura, *J. Mol. Catal. A: Chem.*, 1998, **130**, 1–28.

40 R. S. Downing, P. J. Kunkeler and H. van Bekkum, *Catal. Today*, 1997, **37**, 121–136.

41 C. Boix and M. Poliakoff, *J. Chem. Soc., Perkin Trans. 1*, 1999, 1487–1490.

42 Jenkins, McCullough and Booth, *Ind. Eng. Chem.*, 1930, **31**.

43 A. B. Gamble, J. Garner, C. P. Gordon, S. M. J. O'Conner and P. A. Keller, *Synth. Commun.*, 2007, **37**, 2777–2786.

44 C. A. Dornfeld and S. E. Hazlet, *J. Am. Chem. Soc.*, 1944, **66**, 1781–1782.

45 Y. Liu, Y. Lu, M. Prashad, O. Repic and T. J. Blacklock, *Adv. Synth. Catal.*, 2005, **347**, 217–219.

46 V. Guay and P. Brassard, *J. Heterocycl. Chem.*, 1987, **24**, 1649–1652.

47 M. Nüchter, B. Ondruschka, A. Jungnickel and U. Müller, *J. Phys. Org. Chem.*, 2000, **13**, 579–586.

48 S. Cardinal, P.-A. Paquet-Côté, J. Azelmat, C. Bouchard, D. Grenier and N. Voyer, *Bioorg. Med. Chem.*, 2017, **25**, 2043–2056.

49 Z. Zhao and V. Snieckus, *Org. Lett.*, 2005, **7**, 2523–2526.

50 M. Hassam, A. E. Basson, D. C. Liotta, L. Morris, W. A. L. van Otterlo and S. C. Pelly, *ACS Med. Chem. Lett.*, 2012, **3**, 470–475.

Chapter 5: Conclusion

This project can be divided into four parts, each with different aims. Firstly, we set out to improve the acid stability of a known indole scaffold, as well as probe the Val179 pocket to find an optimum chain length that is best accommodated in the NNRTI-BP. Next, the solvent accessible cleft of the NNRTI-BP was explored by the introduction of various pyridinyl groups in the 2-position of a known active indole NNRTI. The indole base scaffold was also used in an effort to develop an irreversible NNRTI through the introduction of suitable groups in the indole 3-position. Lastly, we aimed to investigate the possibility of the diaryl ether scaffold, previously unexplored by our group, as a superior NNRTI.

In order to eliminate the possibility of acid-mediated degradation of the indole NNRTI, the previously used methyl sulphide group (**14**, *Figure 5.1*)[1] was successfully replaced with an alkyl or alkene chain through a Wittig reaction with the corresponding ketone, followed by reduction to obtain the alkyl derivatives. The chain length was additionally varied in order to determine an optimum length. Biological results revealed that a chain length of two carbon atoms was ideal. In addition, four (**18**, **21**, **22** and **23**) out of the twelve derivatives proved to be superior to the methyl sulphide counterpart (**14**, 0.039 μM) and six (shown in *Figure 5.1*) were more potent than nevirapine (0.091 μM) against wild-type HIV-1.

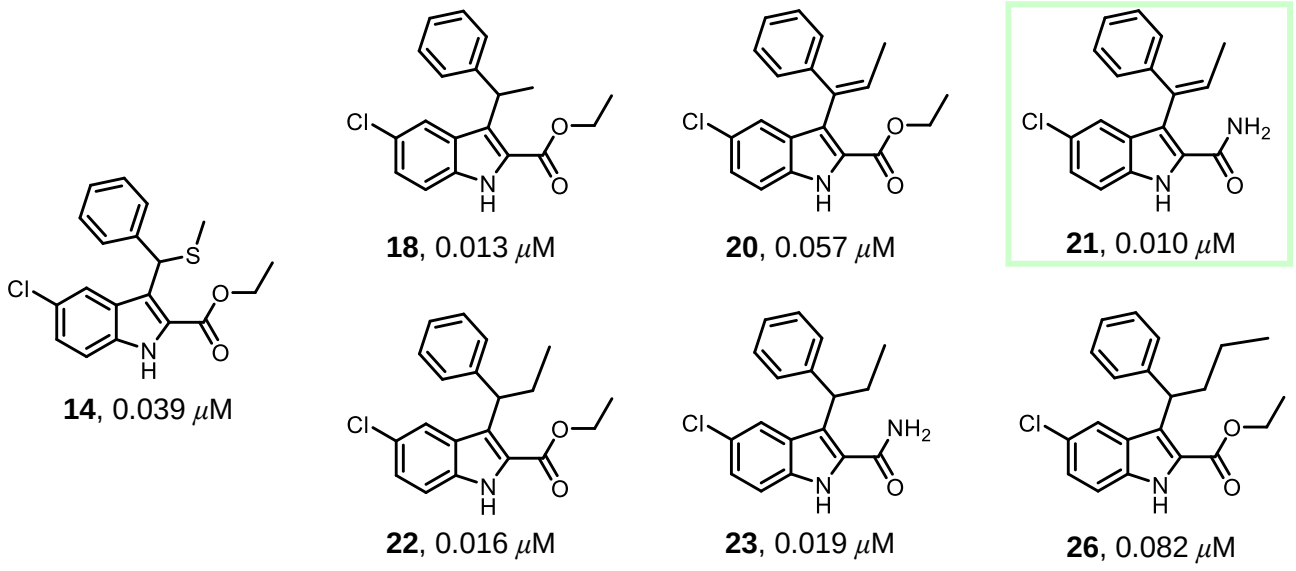


*Figure 5.1: The most active compounds obtained from this section of the work involving minor modifications to a known structure (**14**) to improve the acid stability of this NNRTI as well as probe the Val179 pocket with varying chain lengths in the benzylic position. The phenotypic assay values (IC_{50}) of these derivatives against wild type HIV-1 is shown for each compound in μM.*

A known methyl sulphide scaffold (**15**, *Figure 5.2*) was then successfully modified in the 2-position with various pyridinyl groups introduced using the coupling agent, CDI, the appropriate amine and

the indole 2-carboxylic acid. Three out of the five resultant compounds performed better than the previously synthesised ester counterpart (**15**, 0.060 μM)[1] and the nevirapine control (0.091 μM). In addition, an optimum chain length of one carbon atom and the influence of the pyridinyl nitrogen position was noted. The three most potent compounds (**34** – **36**) were then tested against a panel of mutant HIV viral strains to reveal that compounds **35** and **36** remained potent against HIV-1 viruses with the prevalent NNRTI resistant mutations V106M, Y188C/H, G190A and K103N.

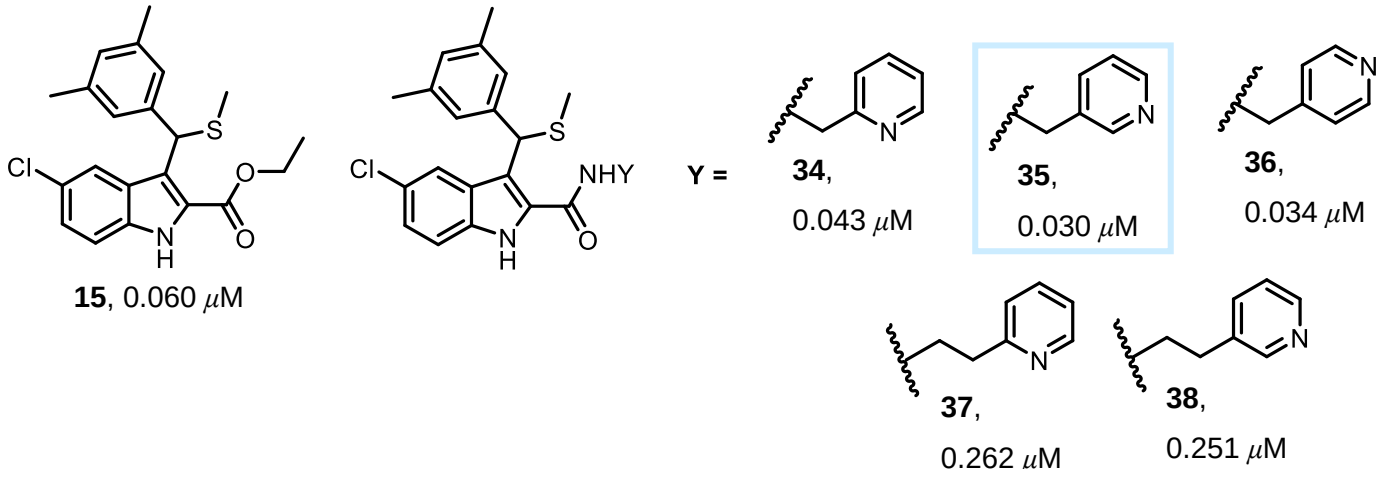


*Figure 5.2: The five final indole carboxamide structures **34** – **38**. The phenotypic assay values (IC_{50}) of these derivatives against wild type HIV-1 is shown for each compound in μM.*

With the indole scaffold performing well as NNRTIs, we designed three proof of concept irreversible indole inhibitors (*Figure 5.3*) to react with the SH group of cysteine in the often present Y181C mutant strain of the virus.[2,3] However, after biological evaluation, these compounds were found to be toxic and inactive against HIV-1.

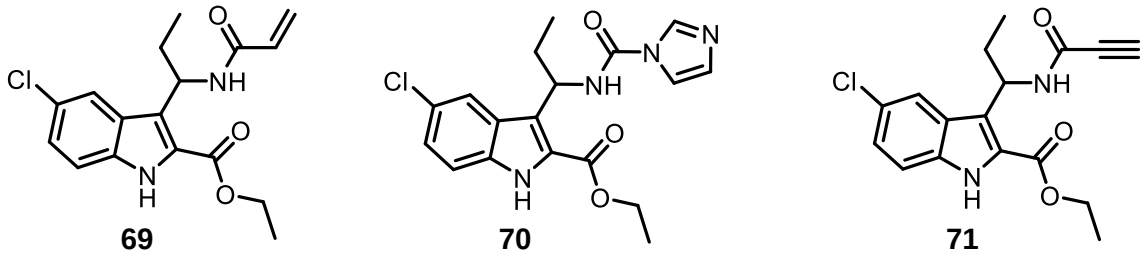
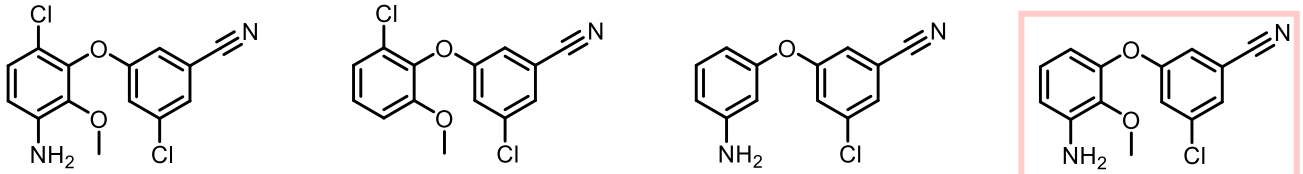


Figure 5.3: Indole NNRTIs with potential as irreversible inhibitors.

The final part of this project involved the design of a synthetic pathway towards a target diaryl ether compound (**100**). This involved an extensive literature study from which a ten-step synthesis was planned.[4,5] Unfortunately, the second to last step in the synthesis, a nucleophilic aromatic substitution, failed and we resorted to synthesising three more easily obtainable proof-of-concept compounds instead. One of these compounds (**118**) proved to be highly potent with an IC_{50} value of 5 nM against wild-type HIV-1 and was in fact the most potent compound synthesised throughout the project. Further biological studies on this compound utilizing a full panel of mutant strains revealed that this compound additionally remained potent against the Y188C mutation.



100, Not obtained **108**, 0.210 μM **112**, 1.122 μM **118**, 0.005 μM

Figure 5.4: Three proof-of-concept diaryl ether NNRTIs with their phenotypic assay values (IC_{50}) against wild-type HIV-1 shown in μM for each compound.

Overall, the project provided valuable SAR data on a known indole scaffold through one minor adjustment, in which a methyl sulphide group was exchanged for an alkyl or alkene chain, and one more significant change, where the larger pyridinyl groups were introduced in the 2-position. The first attempts were also made towards the synthesise of irreversible indole NNRTIs, providing a platform from which to build in the future. From work on our new diaryl ether scaffold, a highly active compound, eighteen times more active than the nevirapine control, resulted. Further optimization, based on the findings of this project, is likely to open research doors and make valuable contributions to the field of antiretroviral drug development.

5.1 Output

Poster presentation December 2016:
P. M. Wessels, W. A. L. van Otterlo, S. C. Pelly, The Design and Synthesis of Possible Non-Nucleoside Reverse-Transcriptase Inhibitors. Presented at the Frank Warren Conference, Rhodes University, Grahamstown, South Africa.

See Section 6.5 for further outputs in progress.

5.2 Bibliography

1 S. Brigg, N. Pribut, A. E. Basson, M. Avgenikos, R. Venter, M. A. Blackie, W. A. L. van Otterlo and S. C. Pelly, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 1580–1584.

2 A. H. Chan, W.-G. Lee, K. A. Spasov, J. A. Cisneros, S. N. Kudalkar, Z. O. Petrova, A. B. Buckingham, K. S. Anderson and W. L. Jorgensen, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 9725–9730.

3 M.-P. de Béthune, *Antiviral Res.*, 2010, **85**, 75–90.

4 L. C. Raiford and D. J. Potter, *J. Am. Chem. Soc.*, 1933, **55**, 1682–1685.

5 R. Hodges and A. Taylor, *J. Am. Chem. Soc.*, 1953, **0**, 4310–4314.

Chapter 6: Future work

6.1 Future work pertaining to Chapter 2

Combining the 2-position of the pyridinyl derivatives (**34** – **38**, *Figure 6.1*) of the indole methyl sulphide known compound **15** with the newly developed acid stable benzylic position (of compound **22** for example, *Figure 6.1*), may yield interesting results. This may be achieved through combining the methods used previously to obtain two separate scaffolds (*Scheme 6.1*). Firstly, the 3,5-dimethylbenzoyl moiety may be introduced in the 3-position of the starting indole, ethyl 5-chloro-1*H*-indole-2-carboxylate (**40**), through a Friedel Crafts acylation reaction with 3,5-dimethylbenzoyl chloride (**49**). The resultant ketone may then be converted into a propenyl group through a Wittig reaction with ethyltriphenylphosphonium bromide, after which reduction with Pd/C should yield the saturated propyl moiety in the benzylic position. The ester in the indole 2-position may then be hydrolysed to the acid, which, through the use of the coupling agent CDI, can be converted into a desired amide.

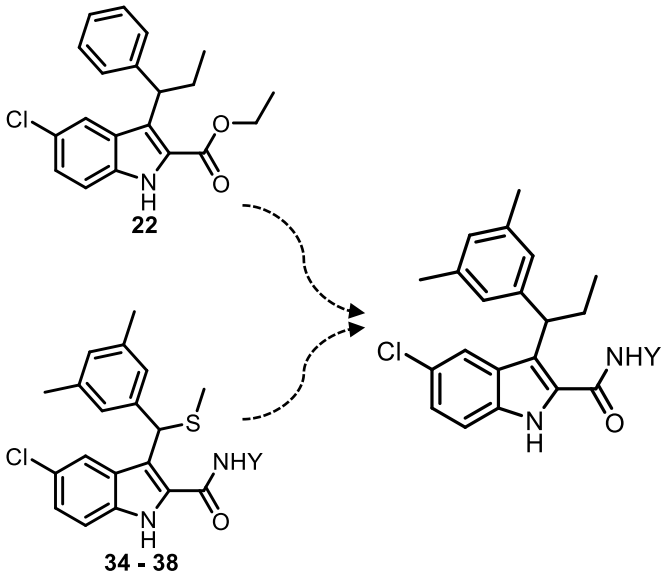
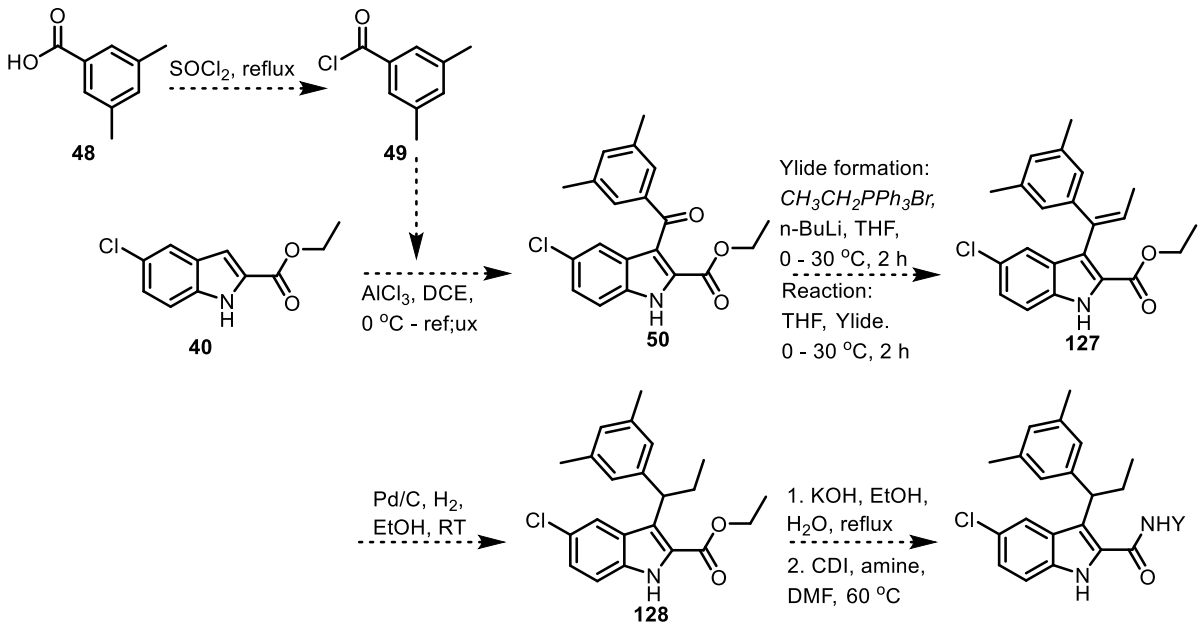


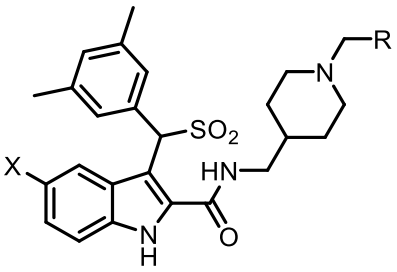
Figure 6.1: A combination of structural features introduced onto the indole scaffold in Chapter 2.



Scheme 6.1: A suggested method to obtain an indole NNRTI with a combination of features from Chapter 2.

With great variety introduced in the indole 2-position over the past years, the effects on activity of many different structures may be explored. Based on recent successes obtained by Li *et al. (Table 6.1),*[1] we propose the introduction of the piperidine structures below (*Table 6.2*) as a springboard for this study.

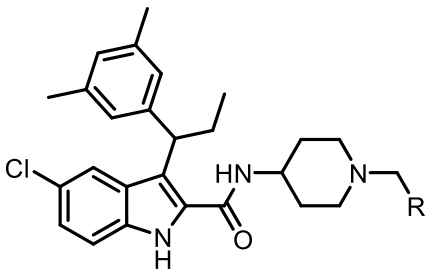
Table 6.1: The most potent compounds synthesised by Li et al. reported in "Discovery of novel piperidine-substituted indolylarylsulfones as potent HIV NNRTIs via structure-guided scaffold morphing and fragment rearrangement" published in the European Journal of Medicinal Chemistry in 2017.[1]



129 X = Br
130 - 133 X = Cl

#	**R**	EC_{50} (μM)								CC_{50} (μM)
		IIIB	L100I	K103N	Y181C	Y188L	E138K	F227L/V106A	RES056	
129	-CH_2OH	0.006 ± 0.002	0.017 ± 0.008	0.084 ± 0.015	0.13 ± 0.011	>6.3	0.14 ± 0.01	0.22 ± 0.15	>6.3	6.3 ± 5.8
130	4-$CONH_2$-Ph	0.012 ± 0.005	0.026 ± 0.002	0.13 ± 0.10	0.29 ± 0.11	>3.2	0.10 ± 0.02	0.30 ± 0.05	>5.2	8.8 ± 5.4
131	Pyridine-3-yl	0.012 ± 0.008	0.034 ± 0.004	0.22 ± 0.06	0.56 ± 0.06	4.8 ± 0.8	0.14 ± 0.10	0.39 ± 0.19	>5.9	6.9 ± 5.2
132	4-methyl-2*H*-1,2,4-triazol-3(4*H*)-one 3-yl	0.009 ± 0.004	0.049 ± 0.007	0.18 ± 0.02	0.20 ± 0.05	0.84 ± 0.21	0.043 ± 0.8	0.27 ± 0.09	>14.0	14.0 ± 4.8
133	4-$CONH_2$-Ph	0.011 ± 0.008	0.029 ± 0.005	0.12 ± 0.01	0.29 ± 0.09	4.7 ± 1.6	0.086 ± 0.00	0.17 ± 0.09	4.4 ± 1.5	6.2 ± 2.1

Table 6.2: Suggested groups for introduction into the indole 2-position.



R
-CH_2OH
4-$CONH_2$-Ph
Pyridine-3-yl
4-methyl-2*H*-1,2,4-triazol-3(4*H*)-one 3-yl
4-$CONH_2$-Ph

6.2 Future work pertaining to Chapter 3

The future work pertaining to this section will rely largely on molecular modelling as a design tool, since the bond-forming functionality may only bind when correctly positioned in the enzyme. Firstly, the non-covalent binding affinity will be considered after which, if it seems promising and the molecule is correctly positioned to make covalent binding possible, the molecule will be docked allowing for covalent bond formation. Protein and ligand preparation and the docking of the molecules will be performed using the Schrödinger Maestro 11 modelling interface. The protein, the HIV-1 RT enzyme, will be obtained from the Protein Data Bank. Visual analysis of the molecule docked in the NNRTI-BP will play a large role before binding energy calculations will be considered.

Some options for additional electrophilic warheads are illustrated in *Figure 6.2*. For the incorporation of the first three bond-forming functionalities (**129** – **131**) into the molecule, we propose the strategy described in Chapter 3 were an amine in the benzylic position will form part of an amide structure in the final compound. This may be achieved through substitution or coupling reactions. A detailed literature study on carbon-carbon bond forming reactions will then be undertaken in order to identify the most suitable methods for the introduction of the other electrophilic groups (**132** – **135**). After the synthesis of these compounds, we hope to be able to obtain crystal structures of the molecules in the enzyme or protein HRMS to confirm covalent bond formation or to make further progress towards the design of such an NNRTI. Trial and error synthesis and working in close collaboration with our biological analyst will form a large part of this study.

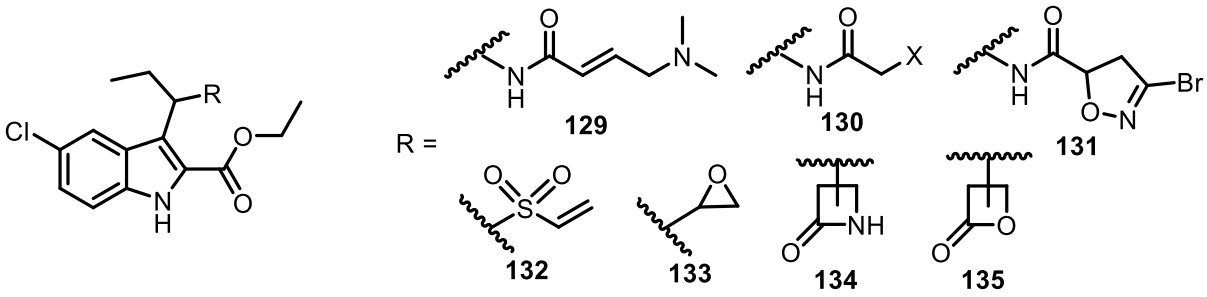


Figure 6.2: Suggestions for an electrophilic warhead that may be incorporated as R-group for covalent binding with residues in the NNRTI-BP.

6.3 Future work pertaining to Chapter 4

As discussed in Chapter 4, Section 4.2.5, for a successful S_NAr reaction, an electron-withdrawing group is usually required *ortho* or *para* to the leaving group to assist with electron stabilization upon attack by the nucleophile.[2] In order to obtain the target compound **100**, which we were unable to synthesise during this project (*Figure 6.3*), we therefore propose switching the fluoro and hydroxy groups on the rings so that the chloro group on compound **136** could act as an electron stabilizing group *ortho* to the leaving group (Option 1a, *Figure 6.3*). If the chloro group is not effective enough

as an electron stabilizer, a group such as a carbonyl, nitro, or cyanide group may be introduced *para* to the leaving group (**138**, Option 1b). Alternatively, the fluoro and hydroxy groups need not be switched if an electron withdrawing group is situated *ortho* or *para* to the fluoro group as in Option 2 (**140** or **142**). The positioning of the cyanide group will, in this case, depend on molecular modelling results.

The starting materials required for option 1a are 1-chloro-2-fluoro-3-methoxy-4-nitrobenzene (**136**) and 3-chloro-5-hydroxybenzonitrile (**137**). 1-Chloro-2-fluoro-3-methoxy-4-nitrobenzene (**136**) may be obtained from 6-chloro-2-methoxy-3-nitrophenol (**97**) through the reaction with a fluorinating reagent (**144**) as shown in *Scheme 6.2* below.[3] 3-Chloro-5-hydroxybenzonitrile (**137**) is commercially available and may be purchased from Sigma Aldrich.

For option 1b (*Figure 6.3*), compound **138** may be synthesised from commercially available 5-chlorovanillin (**86**) following a procedure described by Hodges and Taylor (*Scheme 6.3*).[4] The corresponding oxime **89** will be formed first through reaction of 5-chlorovanillin with hydroxylamine, after which the nitrile **90** may be formed by the addition of acetic anhydride to the oxime and heating at reflux. Hodges and Taylor obtained a single product (**91**) from reaction of compound **90** with fuming nitric acid. We anticipate the formation of compound **145** through a deacetylation reaction using a 4 N sodium hydroxide solution at reflux. Next, the same fluorination reaction used in option 1a (*Scheme 6.2*) is expected to yield compound **138** (*Figure 6.3*). A S_NAr reaction with commercially available 3-chloro-5-hydroxybenzonitrile (**137**), followed by nitro reduction, as described in Chapter 4, should yield target compound **139** (*Figure 6.3*).

For the synthesis of target compounds **141** or **143** (Option 2, *Figure 6.3*), the chloro-fluorobenzonitrile reagent may be purchased from Sigma Aldrich. 6-Chloro-2-methoxy-3-nitrophenol (**97**) may be synthesised as described in Chapter 4. After the S_NAr reaction between **97** and **140** or **142**, nitro reduction should yield **141** or **143**.

With the future work planned we hope to improve upon our methods and yields to obtain more potent compounds against HIV-1 in the future.

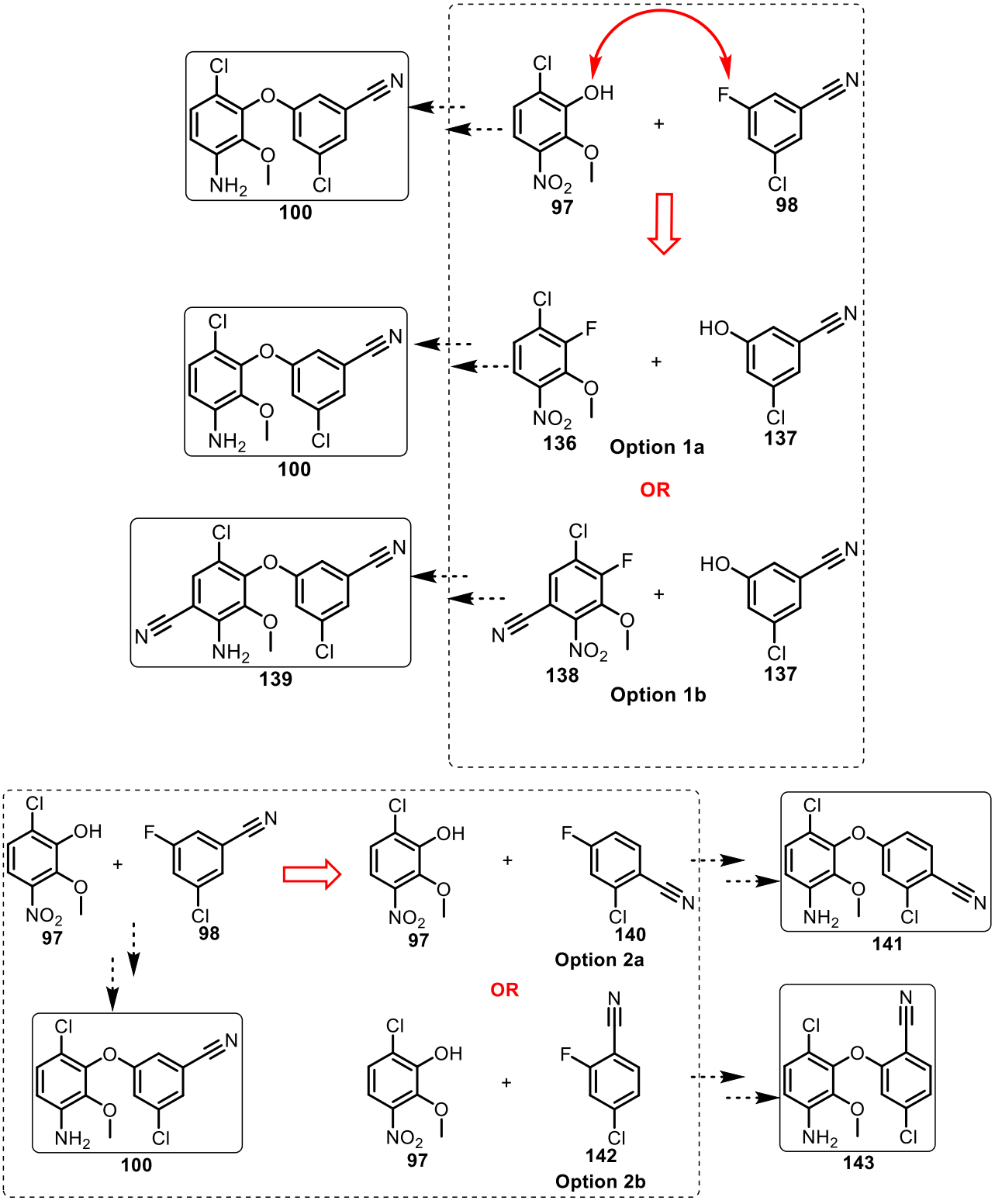
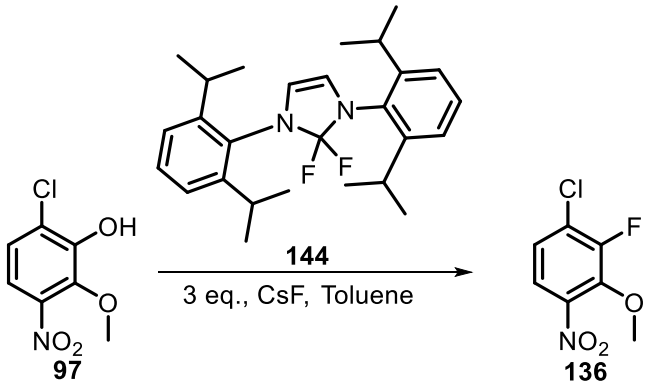
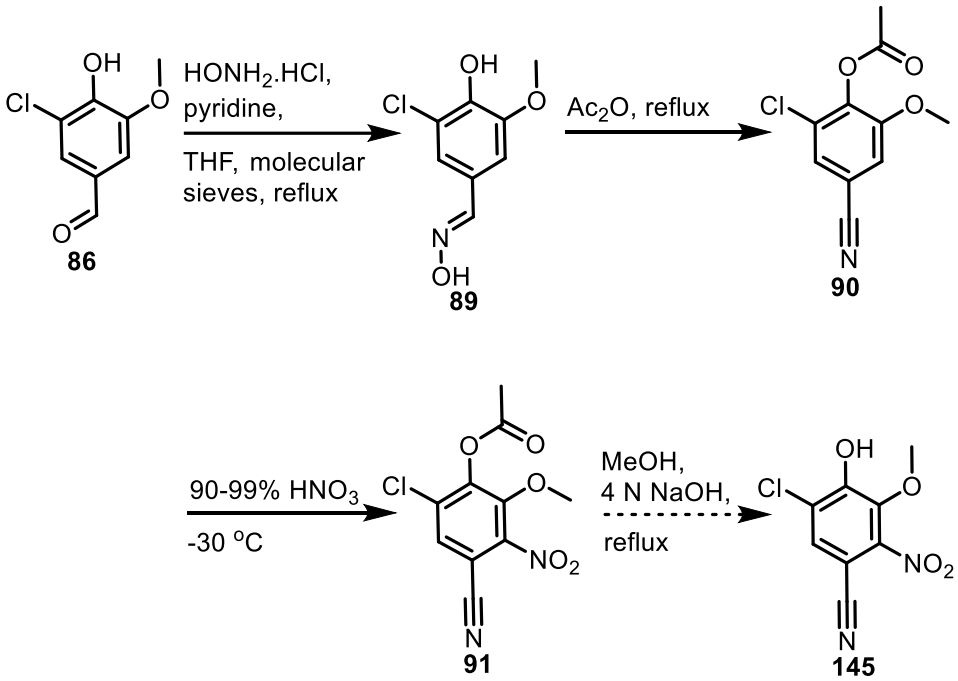


*Figure 6.3: Suggested synthetic pathways to target compound **100** or a similar compound.*



*Scheme 6.2: From OH to F using a fluorinating reagent **144**.*



*Scheme 6.3: A suggested synthetic pathway to compound **145**.*

6.4 Anticipated output

The work reported in this thesis, in particular in Chapter 2, will contribute to the following publication which is currently in progress and will be submitted within the next two to three months:
Title: “Indole based NNRTIs with improved potency against resistant strains of HIV”, Authors: P. M. Wessels, S. Brigg, W. A. L. van Otterlo, S. C. Pelly and A. E. Basson, Journal we aim to publish in is Bioorganic & Medicinal Chemistry Letters. In addition, patenting of these results is currently being considered in discussion with the University’s intellectual property office, INNOVUS.

6.5 Bibliography

1 X. Li, P. Gao, B. Huang, Z. Zhou, Z. Yu, Z. Yuan, H. Liu, C. Pannecouque, D. Daelemans, E. De Clercq, P. Zhan and X. Liu, *Eur. J. Med. Chem.*, 2017, **126**, 190–201.

2 J. Clayden, N. Greevs, S. Warren and P. Wothers, *Organic Chemistry*, Oxford University Press, Oxford, Second., 2001.

3 P. Tang, W. Wang and T. Ritter, *J. Am. Chem. Soc.*, 2011, **133**, 11482–11484.

4 R. Hodges and A. Taylor, *J. Am. Chem. Soc.*, 1953, 4310–4314.